

中外物理学精品书系

经典系列·4

数理物理基础

——物理需用线性高等数学导引

彭桓武 徐锡申 编著



中外物理学精品书系·经典系列4

数理物理基础——物理需用线性高等数学导引

彭桓武（1915–2007），1935年清华大学物理系毕业。1938年冬赴英，1940年和1945年先后获得爱丁堡大学哲学博士和科学博士学位。1947年底回国后历任云南大学、清华大学、北京大学教授，中科院近代物理所、原子能所、高能所研究员、副所长，二机部九院副院长，理论物理所所长，曾多次在北京大学、中国科技大学及其研究生院授课。1955年当选为中科院首批院士。

他早年在英国从事固体理论、量子物理和量子场论的研究。回国后，他领导了我国反应堆和核武器的理论设计工作，在固体和统计物理、原子和分子物理等领域作出了卓越的贡献，培养了一大批优秀的科学工作者。

彭桓武曾荣获国家自然科学奖一等奖(1982)、国家科技进步奖特等奖(1985)、何梁何利基金科学与技术成就奖(1995)，“两弹一星功勋奖章”(1999)。为了表彰彭桓武在理论物理领域上取得的成就，以及他为中国科学研究所作的贡献，2006年6月13日，经国际天文学联合会小天体命名委员会批准，将国际永久编号为第48798号小行星，命名为“彭桓武星”。

徐锡申 1924生。1950年清华大学物理系毕业，1954年北京大学研究生院毕业。1954年至1960年在西北大学物理系工作。1960年初调至九院九所，长期参加和负责国家重点工程物理项目中物质的物态方程和辐射性质等课题的理论研究，有多项科研成果获奖。曾多次讲授“量子统计物理学”的研究生课程，著有《实用物态方程理论导引》，另有译著多种。


ISBN 978-7-301-20026-1
9 787301 200261 >
定价：60.00元



中外物理学精品书系

经典系列 · 4

数理物理基础
——物理需用线性高等数学导引

彭桓武 徐锡申 编著



图书在版编目(CIP)数据

数理物理基础:物理需用线性高等数学导引/彭桓武,徐锡申编著.—北京:北京大学出版社,2012.7

(中外物理学精品书系·经典系列)

ISBN 978-7-301-20026-1

Ⅰ.①数… Ⅱ.①彭… ②徐… Ⅲ.①数学物理方法 Ⅳ.①O411.1

中国版本图书馆 CIP 数据核字(2011)第 281198 号

书　　名:**数理物理基础——物理需用线性高等数学导引**

著作责任者:彭桓武　徐锡申　编著

责 任 编 辑:周月梅　顾卫宇

标 准 书 号:ISBN 978-7-301-20026-1/O·0858

出 版 发 行:北京大学出版社

地　　址:北京市海淀区成府路 205 号　100871

网　　址:http://www.pup.cn

电　　话:邮购部 62752015　发行部 62750672　编辑部 62752021
出版部 62754962

电 子 信 箱:zpup@pup.pku.edu.cn

印　刷　者:北京中科印刷有限公司

经　销　者:新华书店

730 毫米×980 毫米　16 开本　21.25 印张　400 千字

2001 年 5 月第 1 版

2012 年 7 月重排　2012 年 7 月第 3 次印刷

定　　价:60.00 元

《中外物理学精品书系》
编 委 会

序　言

物理学是研究物质、能量以及它们之间相互作用的科学。她不仅是化学、生命、材料、信息、能源和环境等相关学科的基础，同时还是许多新兴学科和交叉学科的前沿。在科技发展日新月异和国际竞争日趋激烈的今天，物理学不仅囿于基础科学和技术应用研究的范畴，而且在社会发展与人类进步的历史进程中发挥着越来越关键的作用。

我们欣喜地看到，改革开放三十多年来，随着中国政治、经济、教育、文化等领域各项事业的持续稳定发展，我国物理学取得了跨越式的进步，做出了很多为世界瞩目的研究成果。今日的中国物理正在经历一个历史上少有的黄金时代。

在我国物理学科快速发展的背景下，近年来物理学相关书籍也呈现百花齐放的良好态势，在知识传承、学术交流、人才培养等方面发挥着无可替代的作用。从另一方面看，尽管国内各出版社相继推出了一些质量很高的物理教材和图书，但系统总结物理学各门类知识和发展，深入浅出地介绍其与现代科学技术之间的渊源，并针对不同层次的读者提供有价值的教材和研究参考，仍是我国科学传播与出版界面临的一个极富挑战性的课题。

为有力推动我国物理学研究、加快相关学科的建设与发展，特别是展现近年来中国物理学者的研究水平和成果，北京大学出版社在国家出版基金的支持下推出了《中外物理学精品书系》，试图对以上难题进行大胆的尝试和探索。该书系编委会集结了数十位来自内地和香港顶尖高校及科研院所的知名专家学者。他们都是目前该领域十分活跃的专家，确保了整套丛书的权威性和前瞻性。

这套书系内容丰富，涵盖面广，可读性强，其中既有对我国传统物理学发展的梳理和总结，也有对正在蓬勃发展的物理学前沿的全面展示；既引进和介绍了世界物理学研究的发展动态，也面向国际主流领域传播中国物理的优秀专著。可以说，《中外物理学精品书系》力图完整呈现近现代世界和中国物理

科学发展的全貌，是一部目前国内为数不多的兼具学术价值和阅读乐趣的经典物理丛书。

《中外物理学精品书系》另一个突出特点是，在把西方物理的精华要义“请进来”的同时，也将我国近现代物理的优秀成果“送出去”。物理学科在世界范围内的重要性不言而喻，引进和翻译世界物理的经典著作和前沿动态，可以满足当前国内物理教学和科研工作的迫切需求。另一方面，改革开放几十年来，我国的物理学研究取得了长足发展，一大批具有较高学术价值的著作相继问世。这套丛书首次将一些中国物理学者的优秀论著以英文版的形式直接推向国际相关研究的主流领域，使世界对中国物理学的过去和现状有更多的深入了解，不仅充分展示出中国物理学研究和积累的“硬实力”，也向世界主动传播我国科技文化领域不断创新的“软实力”，对全面提升中国科学、教育和文化领域的国际形象起到重要的促进作用。

值得一提的是，《中外物理学精品书系》还对中国近现代物理学科的经典著作进行了全面收录。20世纪以来，中国物理界诞生了很多经典作品，但当时大都分散出版，如今很多代表性的作品已经淹没在浩瀚的图书海洋中，读者们对这些论著也都是“只闻其声，未见其真”。该书系的编者们在这方面下了很大工夫，对中国物理学科不同时期、不同分支的经典著作进行了系统的整理和收录。这项工作具有非常重要的学术意义和社会价值，不仅可以很好地保护和传承我国物理学的经典文献，充分发挥其应有的传世育人的作用，更能使广大物理学人和青年学子切身体会我国物理学研究的发展脉络和优良传统，真正领悟到老一辈科学家严谨求实、追求卓越、博大精深的治学之美。

温家宝总理在2006年中国科学技术大会上指出，“加强基础研究是提升国家创新能力、积累智力资本的重要途径，是我国跻身世界科技强国的必要条件”。中国的发展在于创新，而基础研究正是一切创新的根本和源泉。我相信，这套《中外物理学精品书系》的出版，不仅可以使所有热爱和研究物理学的人们从中获取思维的启迪、智力的挑战和阅读的乐趣，也将进一步推动其他相关基础科学更好更快地发展，为我国今后的科技创新和社会进步做出应有的贡献。

《中外物理学精品书系》编委会　主任

中国科学院院士，北京大学教授

王恩哥

2010年5月于燕园

内 容 提 要

本书简明扼要地阐述了解决物理问题需用的重要数学概念、方法和定理及其运用条件；从线性变换入手，提纲挈领地将线性数学的有关内容连贯起来，构成有机整体. 内容包括：线性变换，群，行列式，线性方程组的求解，矢量与张量分析，二次型和主轴变换，线性积分方程，函数空间，变分法，微分方程绪论，二阶线性偏微分方程，二阶线性常微分方程，微分方程的数值解法等共十三章.

本书可供高等院校物理系及其他相关专业师生作为参考教材，亦可供广大有关科技工作者自学参考之用.

序

1950年秋，本书第一作者(彭)在清华大学物理系，首次开一门选修的数学课，内容涉及多门数学，但都是学物理者经常需要懂和用的那些部分．因为当时是新课，听课同学包括大学高年级生和研究生．每周两课时，历时两学期，无课本，无讲义，仅在上学期曾经陆续记下简单的授课大纲，累计满一本薄薄的练习本．学年结束时以检察听课笔记代替考试．笔记最好的为研究生徐锡申和四年级生钱尚武所记，两人乃被邀誊抄其全部笔记与彭保存．本书即根据上述资料，由第二作者(徐)整理和稍加补充而成．

彭在吉林毓文中学读初中三年级时，所用数学课本即为三角、几何、小代数的混合初等数学教材，而彭在清华大学物理系所开的数学课，则好像是代数、几何、分析的混合高等数学．根据彭的科研经历，到他开数学课时，他主要注意在物理的量子理论方面，有量子态的叠加原理作根据．碰到的数学问题，管它是代数方程、微分积分方程，总是线性的；这就使这部分高等数学便于混合地讲，似乎线性高等数学也有个统一的基础．

彭当时用"数理物理"作为课程名称，授课时主要阐明重要数学概念、方法和定理及其条件，旨在为学物理者经常需要弄懂和运用的数学作多方面的简要引导．因此本书内容与学习苏联后的"数学物理方法"教学大纲差别较大；也与风格各异的德、英、美等的几家"数理物理"或"理论物理方法"专著，博浅专深有所不同．所以本书书名采用"数理物理基础"．

徐在整理书稿过程中，为了表述的准确与充实，除查用过大部分原引参考书之外，还参阅了以后直至近年来出版的多种书籍并有所引用．另外，对全书章节进行详细编排后，适当补充了少量内容．最后，经过彭细致审阅校改后定稿．具体内容包括：线性变换，群，行列式，线性方程组的求解，矢量与张量分析，二次型和主轴变换，线性积分方程，函数空间，变分法，微分方程绪论，二阶线性偏微分方程，二阶线性常微分方程，微分方程的数值解法等共十三章．

本书具有以下特点：紧密联系物理问题，重点阐述重要数学概念和方法．

从线性变换入手，提纲挈领地将线性数学的有关内容连贯起来，构成有机整体. 从典型特例着手，深入浅出，循序渐进，从而引出一般理论，易于学习. 叙述力求严格，繁冗论证有时借助物理直观，强调运用条件，便于正确运用. 关注到数值方法的重要性，多处有着重论述. 阐述简明扼要，以极少篇幅而涵盖相当丰富内容，可快速引导入门. 尽管跨过了半个世纪，但现在看来，本书的选材内容和讲授方法，国内外有关书籍仍极少见.

最后，对于钱尚武教授提供的听课笔记，表示衷心感谢. 还要感谢清华大学张泽瑜教授，他曾关怀过本书的最初出版事宜. 作者对《北京大学物理学丛书》编委们致谢，特别对常务编委周月梅女士和北京大学出版社有关人士提供的方便致谢.

谨以此书作为向新世纪暨母校清华大学九十周年校庆的献礼.

彭桓武 中国科学院理论物理研究所

徐锡申 北京应用物理与计算数学研究所

2001 年 3 月于北京

目　　录

第1章　线性变换

我们想首先引进代数和几何中经常碰到的一些基本概念.

1.1　线性变换的定义①

通过系数 a_{ik} 为给定的一组 n 个**线性方程**

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1, \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2, \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= y_n,
\end{aligned}
$$

或者简写为

$$\sum_{k=1}^{n} a_{ik}x_k = y_i \quad (i = 1, \cdots, n),$$

则对于每一个值组$(x_1, x_2, \cdots, x_n)$将有唯一的值组$(y_1, y_2, \cdots, y_n)$与之相对应. 我们将这种对应称之为从值组$(x_1, x_2, \cdots, x_n)$到值组$(y_1, y_2, \cdots, y_n)$的一个**线性变换**.

备注　(i) 一个线性变换完全由特定系数值 a_{ik} 所决定.

(ii) 从值组$(x_1, x_2, \cdots, x_n)$到值组$(y_1, y_2, \cdots, y_n)$的这种对应是唯一的. 但是，对于给定值组$(y_1, y_2, \cdots, y_n)$的逆变换问题，或线性方程组的求解问题，则未必如此.

(iii) 这些数值可能是复数(**C**)或实数(**R**)，也有可能是形成一个**域 F** 的任何其他数.(关于域的概念见本节附录.)

关于备注(ii)前一句的证明：只要 a_{ik} 与$(x_1, x_2, \cdots, x_n)$是域中的元素，则根据域的定义即可证明$(y_1, y_2, \cdots, y_n)$是唯一的.

关于备注(ii)后一句的内容将在本章 1.6 节以及第 4 章(线性方程组的求解)中予以讨论.

---

① R. Courant, D. Hilbert, *Methoden der Mathematische Physik*, Bd. Ⅰ, Springer, Berlin, 1930, S. 5.[中译本：R. 柯朗, D. 希尔伯特著,《数学物理方法》,卷Ⅰ,钱敏,郭敦仁译,科学出版社,北京,1958,4 页.]

1.1A 附录：域的概念[①②]

域 F 是任何这样的元素集合，对于集合中元素定义有加法与乘法的运算，它们具有下列性质：

(i) 对于 F 中任何一对元素 λ',λ''，其和 $\lambda'+\lambda''$ 与其积 $\lambda'\lambda''$ 是唯一确定的，且同属 F 中元素. 加法和乘法都满足

(a) 交换律：

$$\lambda'+\lambda''=\lambda''+\lambda',\quad \lambda'\lambda''=\lambda''\lambda';$$

(b) 结合律：

$$\lambda'+(\lambda''+\lambda''')=(\lambda'+\lambda'')+\lambda''',\quad \lambda'(\lambda''\lambda''')=(\lambda'\lambda'')\lambda''';$$

以及(c)乘法对于加法的分配律：

$$\lambda'(\lambda''+\lambda''')=\lambda'\lambda''+\lambda'\lambda'''.$$

(ii) F 中含有两个独特的元素：0(零元素)与 1(单位元素)，使得对 F 中每一元素 λ 都有

$$\lambda+0=\lambda,\quad 1\cdot\lambda=\lambda.$$

(iii) 对于 F 中的每个元素 λ，$x+\lambda=0$ 在 F 中有一个解，称为元素 λ 的负元素，并用 $-\lambda$ 表示.

(iv) 对于 F 中的每个元素 $\lambda\neq0$，$x\lambda=1$ 在 F 中有一个解，称为元素 λ 的逆元素，并用 λ^{-1} 表示.

或者简要地说：一个集合 F，其中任意二元素之和与积亦含于 F 中，加法与乘法满足交换律、结合律与分配律；对于加法，F 有零元素，每个元素有负元素；对于乘法，F 有单位元素，除去零元素外，每个元素有逆元素；这样的代数结构就称为**域**.

有了实际的东西，才会产生抽象的理论，一个概念必定包含许多内容，数学家所以引进域就是因为在数学中有很多集合都满足上述条件. 例如：

Q(有理数域)　{全体有理数}，

R(实数域)　{全体实数}，

C(复数域)　{全体复数}，

Q($\sqrt{2}$)　{全体形如 $a+b\sqrt{2}$ 的数，其中 a,b 是有理数}.

或者比 **Q**($\sqrt{2}$) 较一般地，如

Q($\sqrt{m}$)　{全体形如 $a+b\sqrt{m}$ 的数，其中 a,b 是有理数，而 m 可为 2，3，5，6，7，…，只要 m 不是无限即可}

[$m=1,4,9,\cdots$ 也可，不过这样有 **Q**($\sqrt{m}$) 包含在 **Q** 中].

---

① G. Birkhoff, S. MacLane, *A Survey of Modern Algebra*, MacMillan, New York, 1941.

② 参考：《中国大百科全书：数学》，818 页，域(丁石孙撰)，中国大百科全书出版社，北京，1988.

1.2　线性变换的矩阵表示

对于进一步的讨论，方便的是引进矩阵记法；它不仅对缩写方程很有用，而且同时它还代表比数具有更高度复杂性的一个数学对象.

1.2.1　矩阵的定义

一组 $m\times n$ 个数(域中元素)以一定方式有序地排列成的下述阵列

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

称为一个 $m\times n$ **矩阵** A 或$\{a_{ik}\}$，其中 i 为行标，$i=1,2,\cdots,m$；k 为列标，$k=1,2,\cdots,n$.

一个 $m\times n$ 矩阵可以看成是有 n 个未知数 $x_k(k=1,\cdots,n)$的 m 个线性方程的系数有序排列而成：

$$\sum_{k=1}^{n} a_{ik}x_k = y_i \quad (i=1,2,\cdots,m).$$

所以，矩阵与线性变换有着密切的关系.

如果行的数目 m 与列的数目 n 相等，即 $m=n$，则矩阵 A 称为 n **阶矩阵**或 n **阶方阵**.

如果仅有一行(或一列)，则称为一个**行矢量**(或**列矢量**)；简言之，一个**矢量**. 例如：行矢量($m=1$，$1\times n$ 矩阵)

$$\{a_{1k}\} = [a_{11} \quad a_{12} \quad \cdots \quad a_{1n}],$$

列矢量($n=1$，$m\times 1$ 矩阵)

$$\{a_{i1}\} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}.$$

仅有一行和一列的矩阵，即 1×1 矩阵，称为**标量**.

1.2.2　矩阵的运算

下述关于**矩阵的运算**是有用的(**矢量的运算**作为特例包括在内).

(1) 矩阵的加法

两个 $m\times n$ 矩阵 A 与 B 可以相加,其和为一 $m\times n$ 矩阵 C,则有

$$A+B=\{a_{ik}\}+\{b_{ik}\}=\{a_{ik}+b_{ik}\}=\{c_{ik}\}=C,$$

即

$$c_{ik}=a_{ik}+b_{ik}.$$

显然,矩阵的加法满足交换律:

$$A+B=B+A.$$

(2) 矩阵的数乘

矩阵 A 和数 λ 可以相乘,结果为

$$\lambda A=\lambda\{a_{ik}\}=\{\lambda a_{ik}\}=\{a_{ik}\}\lambda=A\lambda;$$

可见,数乘满足交换律.

(3) 矩阵的相等

两个 $m\times n$ 矩阵 A 与 B 相等必须其对应元素都相等,即

$$A=B \quad 意味着 \quad a_{ik}=b_{ik};$$

于是

$$A-B=\{a_{ik}\}-\{b_{ik}\}=\{a_{ik}\}+\{-b_{ik}\}$$
$$=\{a_{ik}-b_{ik}\}=\{0_{ik}\}=0,$$

其中 $0=\{0_{ik}\}$ 为**零矩阵**,具有 $0_{ik}=0$.

若 $A=\{a_{ik}\}$ 为已知矩阵,$X=\{x_{ik}\}$ 为未知矩阵,则

$$X+A=0$$

有解为

$$X=-A,$$

$-A$ 为矩阵 A 的**负矩阵**. 因为 $X+A=0$ 相当于

$$\{x_{ik}\}+\{a_{ik}\}=\{x_{ik}+a_{ik}\}=\{0_{ik}\},$$

即

$$x_{ik}+a_{ik}=0_{ik}=0 \quad 或 \quad x_{ik}=-a_{ik},$$

故有

$$X=\{x_{ik}\}=\{-a_{ik}\}=-\{a_{ik}\}=-A.$$

(4) 矩阵的[链式]乘法

一个 $m\times r$ 矩阵 $A=\{a_{ij}\}$ 和一个 $r\times n$ 矩阵 $B=\{b_{jk}\}$ 可以相乘,其链积 AB 是一个 $m\times n$ 矩阵 $C=\{c_{ik}\}$($i=1,2,\cdots,m,j=1,2,\cdots,r,k=1,2,\cdots,n$),即

$$AB=C, \quad 带有 \quad c_{ik}=\sum_{j=1}^{r}a_{ij}b_{jk}.$$

备注 (i) 左因子 A 的列数 r 必须与右因子 B 的行数 r 相同,因而有**链式乘法(链乘)**之名.

(ii) 对 c_{ik} 适用**"行乘列"**法则.

(iii) 一般来说,矩阵乘法不满足交换律,即 $AB\neq BA$. 因为,若 $m\neq n$,则由备注(i)知链积 BA 不存在;而若 $m=n$,但这时 AB 为 n 阶方阵而 BA 则为 r 阶方阵,也

不可能相等;即使 A,B 为同阶方阵,由于"行乘列"法则(备注(ii)),一般也会使得 $AB \neq BA$.

(5) 矩阵乘法满足结合律

$$(\lambda\mu)A = \lambda(\mu A), \quad (\lambda A)B = \lambda(AB),$$
$$A(BC) = (AB)C.$$

关于 $A(BC)=(AB)C$ 的证明如下:令 A 为 $m \times r$ 矩阵,B 为 $r \times s$ 矩阵,C 为 $s \times n$矩阵,则有

$$\{A(BC)\}_{ik} = \sum_{j=1}^{r} a_{ij}(BC)_{jk} = \sum_{j=1}^{r} a_{ij} \sum_{l=1}^{s} b_{jl} c_{lk},$$
$$\{(AB)C\}_{ik} = \sum_{l=1}^{s} (AB)_{il} c_{lk} = \sum_{l=1}^{s} \sum_{j=1}^{r} a_{ij} b_{jl} c_{lk},$$

如果 r 与 s 为有限则两者相等,因而得证.

(6) 矩阵运算满足分配律

$$(\lambda + \mu)A = \lambda A + \mu A,$$
$$\lambda(A + B) = \lambda A + \lambda B,$$
$$(A + B)C = AC + BC,$$
$$A(B + C) = AB + AC;$$

等等. 以上所有结果都可根据矩阵的加法与乘法的定义,以及矩阵元 a_{ik},…与 λ,μ 等均是域中元素而得以证明.

1.2.3 线性变换的矩阵表示

有了矩阵记法,就可以将**线性变换**用

$$A\boldsymbol{x} = \boldsymbol{y}$$

来描述,其中 A 为给定的**系数矩阵**,而 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 为[**列**]**矢量**,即

$$A = \{a_{ik}\} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix};$$

根据矩阵乘法的定义很容易证明上述结果,因为

$$A\boldsymbol{x} = \boldsymbol{y} \quad \text{即} \quad \sum_{k=1}^{n} a_{ik} x_k = y_i \quad (i = 1,2,\cdots,n).$$

于是,可以认为线性变换 A 将每个矢量 $\boldsymbol{x}$ 变换为一个唯一的对应矢量 $\boldsymbol{y}$,即

$$\boldsymbol{x} \to \boldsymbol{y} = A\boldsymbol{x}.$$

1.3　线性变换与线性空间

1.3.1　线性变换的性质

建立起下列对应

$$x \to y = Ax,\quad x' \to y' = Ax',\quad x'' \to y'' = Ax'',\quad \cdots,$$

其线性变换 A 具有下述基本性质：

(i) **齐性**：$\boldsymbol{x} \to \boldsymbol{y}$　蕴涵　$\lambda \boldsymbol{x} \to \lambda \boldsymbol{y}$；

(ii) **加性**：$\boldsymbol{x}' \to \boldsymbol{y}', \boldsymbol{x}'' \to \boldsymbol{y}''$　蕴涵　$\boldsymbol{x}' + \boldsymbol{x}'' \to \boldsymbol{y}' + \boldsymbol{y}''$；

(iii) 更一般地，结合起来称为**线性**：$\boldsymbol{x}' \to \boldsymbol{y}', \boldsymbol{x}'' \to \boldsymbol{y}'', \cdots$，蕴涵

$$\lambda' \boldsymbol{x}' + \lambda'' \boldsymbol{x}'' + \cdots \to \lambda' \boldsymbol{y}' + \lambda'' \boldsymbol{y}'' + \cdots.$$

最后一个性质可证明如下：

$$\begin{aligned}\lambda' \boldsymbol{y}' + \lambda'' \boldsymbol{y}'' + \cdots &= \lambda' A \boldsymbol{x}' + \lambda'' A \boldsymbol{x}'' + \cdots \\ &= A\lambda' \boldsymbol{x}' + A\lambda'' \boldsymbol{x}'' + \cdots \\ &= A(\lambda' \boldsymbol{x}' + \lambda'' \boldsymbol{x}'' + \cdots).\end{aligned}$$

显然，性质(i)和(ii)可作为(iii)的特例得到证明.

从线性变换的上述性质得到启发，可以引进另一几何概念——矢量空间——以使我们的直观的运用更为便利.

1.3.2　矢量空间和矢量子空间①

(1) 矢量空间的定义

矢量空间（常称为**线性空间**）是一堆元素（称为矢量）的集合，它包含(i)任何二矢量 $\boldsymbol{x}', \boldsymbol{x}''$ 之和 $\boldsymbol{x}' + \boldsymbol{x}''$，以及(ii)任何矢量 $\boldsymbol{x}$ 乘以任何标量 λ 后所得之积 $\lambda \boldsymbol{x}$.

例如，所有 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$，每个 x_i 都可取遍域 **F**(复数域 **C** 或实数域 **R**)中元素. 这个空间将用 $\mathbf{V}_n(\mathbf{C})$ 或 $\mathbf{V}_n(\mathbf{R})$ 表示.

(2) 矢量子空间的定义

矢量空间 **V** 的**矢量子空间**（又称**线性子空间**，或简称**子空间**）是 **V** 的子集，它本身相对于 **V** 中的加法和与标量之乘法的运算是一个矢量空间.

例如，对于 $\mathbf{V}_n$ 中任何给定矢量 $\boldsymbol{x}', \boldsymbol{x}'', \cdots, \boldsymbol{x}^{(m)}$ 的所有线性组合

$$\lambda' \boldsymbol{x}' + \lambda'' \boldsymbol{x}'' + \cdots + \lambda^{(m)} \boldsymbol{x}^{(m)}$$

（每个标量 λ 取遍 **C** 或 **R** 中元素）形成 $\mathbf{V}_n$ 的一个矢量子空间，它被称为由 $\boldsymbol{x}', \boldsymbol{x}'', \cdots, \boldsymbol{x}^{(m)}$ 生成的子空间.

---

① G. Birkhoff, S. MacLane, *A Survey of Modern Algebra*, pp. 167—168.

有了矢量空间和矢量子空间的概念，因而有下一小节的定理.

1.3.3 线性变换与矢量空间映射的定理

定理 一个线性变换将一个矢量空间映射到另一个矢量空间.

利用线性变换的线性性质，即加性与齐性的性质(见 1.3.1 节)，这个定理得以证明.

利用矢量子空间的定义，可将上面的定理予以明晰化为下述形式：

定理 线性变换 A 将矢量空间 $\mathbf{V}_n$ 映射到由矢量

$$\boldsymbol{a}' = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad \boldsymbol{a}'' = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \quad \cdots, \quad \boldsymbol{a}^{(n)} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}$$

所生成的矢量子空间 $\mathbf{S}$.

证 可以认为 $\mathbf{V}_n$ 是由矢量

$$\boldsymbol{\xi}' = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{\xi}'' = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \boldsymbol{\xi}^{(n)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

所生成，因为

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{\xi}' + x_2\boldsymbol{\xi}'' + \cdots + x_n\boldsymbol{\xi}^{(n)}.$$

由于

$$A = \{a_{ik}\} = [\boldsymbol{a}', \boldsymbol{a}'', \cdots, \boldsymbol{a}^{(n)}],$$

而显然有

$$\boldsymbol{\xi}' \to A\boldsymbol{\xi}' = \boldsymbol{a}', \quad \boldsymbol{\xi}'' \to A\boldsymbol{\xi}'' = \boldsymbol{a}'', \quad \cdots, \quad \boldsymbol{\xi}^{(n)} \to A\boldsymbol{\xi}^{(n)} = \boldsymbol{a}^{(n)};$$

因此根据 1.3.1 节线性变换的性质(3)，容易得出

$$\boldsymbol{x} = x_1\boldsymbol{\xi}' + x_2\boldsymbol{\xi}'' + \cdots + x_n\boldsymbol{\xi}^{(n)} \to$$
$$\boldsymbol{y} = A\boldsymbol{x} = x_1\boldsymbol{a}' + x_2\boldsymbol{a}'' + \cdots + x_n\boldsymbol{a}^{(n)},$$

即，$\boldsymbol{y}$ 是 $\boldsymbol{a}', \boldsymbol{a}'', \cdots, \boldsymbol{a}^{(n)}$ 的线性组合；因此根据 1.3.2 节之(2)的定义，$\mathbf{S}$ 是由 $\boldsymbol{a}', \boldsymbol{a}'', \cdots, \boldsymbol{a}^{(n)}$ 所生成的矢量子空间. 证毕.

通过下节将要引进的一些概念，上述结果还可得到更进一步的明晰化.

1.4　矢量空间的基

1.4.1　矢量的线性无关与线性相关[①]

定义　矢量 $\boldsymbol{x}', \boldsymbol{x}'', \cdots, \boldsymbol{x}^{(m)}$ 是**线性无关**的，当且仅当关系式

$$\lambda'\boldsymbol{x}' + \lambda''\boldsymbol{x}'' + \cdots + \lambda^{(m)}\boldsymbol{x}^{(m)} = \boldsymbol{0}$$

只有在 $\lambda' = \lambda'' = \cdots = \lambda^{(m)} = 0$ 时才会成立的情况. 否则，矢量 $\boldsymbol{x}', \cdots, \boldsymbol{x}^{(m)}$ 是**线性相关**的.

(1) 例如，在 $\mathbf{V}_n$ 中，矢量

$$\boldsymbol{\xi}' = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{\xi}'' = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \boldsymbol{\xi}^{(n)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

是线性无关的.

证

$$\lambda'\boldsymbol{\xi}' + \lambda''\boldsymbol{\xi}'' + \cdots + \lambda^{(n)}\boldsymbol{\xi}^{(n)} = \begin{bmatrix} \lambda' \\ \lambda'' \\ \vdots \\ \lambda^{(n)} \end{bmatrix} = \boldsymbol{0},$$

根据 1.2.2 节之(3)，它蕴含 $\lambda' = \lambda'' = \cdots = \lambda^{(n)} = 0$. 证毕.

(2) 例如，一个非零矢量是线性无关的，因为 $\lambda\boldsymbol{x} = \boldsymbol{0}$ 蕴含 $\lambda = 0$；但是一个零矢量 $\boldsymbol{0}$ 是线性相关的，因为对于任何 λ 都有 $\lambda\boldsymbol{0} = \boldsymbol{0}$.

通过选择和限定定义中的一些标量恒等于零，上述定义导致下面的定理.

(3) **定理**　一个线性无关矢量集的任何子集是线性无关的.

线性相关的名词来源于下述定理.

(4) **定理**[②]　$\mathbf{V}_n$ 中非零矢量 $\boldsymbol{x}', \boldsymbol{x}'', \cdots, \boldsymbol{x}^{(m)}$ 是线性相关的，当且仅当有一矢量 $\boldsymbol{x}^{(k)}$ 是其余矢量的线性组合.

证　若 $\boldsymbol{x}^{(k)} = \boldsymbol{x}^{(\neq k)}$ 的线性组合，则移项后有

$$\lambda'\boldsymbol{x}' + \lambda''\boldsymbol{x}'' + \cdots + \lambda^{(m)}\boldsymbol{x}^{(m)} = \boldsymbol{0} \text{ 对 } \lambda^{(k)} \neq 0,$$

即 $\boldsymbol{x}', \cdots, \boldsymbol{x}^{(m)}$ 是线性相关的；反过来，若 $\lambda'\boldsymbol{x}' + \cdots + \lambda^{(m)}\boldsymbol{x}^{(m)} = \boldsymbol{0}$，具有比如 $\lambda^{(k)} \neq 0$，

---

① 类似定义也适用于 $m \times n$ 矩阵.

② G. Birkhoff, S. MacLane, *A Survey of Modern Algebra*, pp. 172—173, Theorem 3.

则可解出 $\boldsymbol{x}^{(k)}$ 为

$$\boldsymbol{x}^{(k)} = \boldsymbol{x}^{(\neq k)} \text{ 的线性组合}.$$

证毕.

矢量的线性无关或线性相关这个概念为矢量空间的几何中所固有，通过下面的分析可以使之更加明晰.

1.4.2 矢量空间的基与维数

(1) 基的定义

若一矢量空间可由一线性无关的矢量子集生成，则称此矢量子集为该**矢量空间的一个基**.

例如，

$$\boldsymbol{\xi}' = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \boldsymbol{\xi}^{(n)} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

是矢量空间 $\mathbf{V}_n$ 的一个基，因为 $\mathbf{V}_n$ 中每个矢量

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{\xi}' + x_2\boldsymbol{\xi}'' + \cdots + x^{(n)}\boldsymbol{\xi}^{(n)}$$

均可由 $\boldsymbol{\xi}', \boldsymbol{\xi}'', \cdots, \boldsymbol{\xi}^{(n)}$ 予以生成.

(2) 基的选择

矢量空间的基的选择，很大程度上是任意的. 实际上，我们有下述定理.

定理 $\mathbf{V}_n$ 中的任何线性无关矢量集合 $\boldsymbol{\eta}', \boldsymbol{\eta}'', \cdots, \boldsymbol{\eta}^{(p)}$ ($p \leqslant n$) 可形成 $\mathbf{V}_n$ 的基的一部分.

证 第一步 先证明 $\boldsymbol{\eta}^{(p)}, \boldsymbol{\xi}', \cdots, \boldsymbol{\xi}^{(n-1)}$ 形成 $\mathbf{V}_n$ 的一个基.

因为 $\boldsymbol{\xi}', \cdots, \boldsymbol{\xi}^{(n)}$ 形成 $\mathbf{V}_n$ 的一个基，$\boldsymbol{\eta}^{(p)}$ 可表达为 $\boldsymbol{\xi}', \cdots, \boldsymbol{\xi}^{(n)}$ 的线性组合. 根据 1.4.1 节，矢量 $\boldsymbol{\eta}^{(p)}, \boldsymbol{\xi}', \cdots, \boldsymbol{\xi}^{(n)}$ 是线性相关的，所以

$$\mu^{(p)}\boldsymbol{\eta}^{(p)} + \lambda'\boldsymbol{\xi}' + \cdots + \lambda^{(n)}\boldsymbol{\xi}^{(n)} = \mathbf{0}$$

具有 $\mu^{(p)}, \lambda', \cdots, \lambda^{(n)}$ 不全为零. 这里 $\mu^{(p)} \neq 0$ 以及 $\lambda', \cdots, \lambda^{(n)}$ 不全为零($\mu^{(p)}\eta^{(p)} \neq \mathbf{0}$, $\lambda'\xi' + \cdots + \lambda^{(n)}\xi^{(n)} \neq \mathbf{0}$). 因而，某个 λ(比如 $\lambda^{(n)}$) $\neq 0$ 和因为 $\mu^{(p)} \neq 0$，故 $\boldsymbol{\xi}^{(n)} = \boldsymbol{\eta}^{(p)}, \boldsymbol{\xi}', \cdots, \boldsymbol{\xi}^{(n-1)}$ 的线性组合，肯定包含 $\boldsymbol{\eta}^{(p)}$；所以，(i) 集合 $\boldsymbol{\eta}^{(p)}, \boldsymbol{\xi}', \cdots, \boldsymbol{\xi}^{(n-1)}$ 是线性无关的(如果不然，$\boldsymbol{\eta}^{(p)}$ 可表达为 $\boldsymbol{\xi}', \cdots, \boldsymbol{\xi}^{(n-1)}$ 的线性组合，它与 $\boldsymbol{\xi}^{(n)} = \boldsymbol{\eta}^{(p)}, \boldsymbol{\xi}', \cdots, \boldsymbol{\xi}^{(n-1)}$ 的线性组合两者结合一起会导致与 $\boldsymbol{\xi}', \cdots, \boldsymbol{\xi}^{(n)}$ 的线性无关相矛盾的结果)，以

及(ii) $\mathbf{V}_n$ 中所有矢量,通过 $\boldsymbol{\xi}^{(n)}=\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-1)}$ 的线性组合这个表达式,可用 $\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-1)}$ 来表达,即

$$\boldsymbol{x}=\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-1)}\text{ 的线性组合.}$$

换句话说,$\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-1)}$ 形成 $\mathbf{V}_n$ 的一个基.

第二步　再证明 $\boldsymbol{\eta}^{(p-1)},\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-2)}$ 形成 $\mathbf{V}_n$ 的一个基.

因为 $\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-1)}$ 形成 $\mathbf{V}_n$ 的一个基,$\boldsymbol{\eta}^{(p-1)}$ 可表达为 $\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-1)}$ 的线性组合,于是,$\boldsymbol{\eta}^{(p-1)},\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-1)}$ 是线性相关的,即

$$\mu^{(p-1)}\boldsymbol{\eta}^{(p-1)}+\mu^{(p)}\boldsymbol{\eta}^{(p)}+\lambda'\boldsymbol{\xi}'+\cdots+\lambda^{(n-1)}\boldsymbol{\xi}^{(n-1)}=\mathbf{0},$$

具有 $\mu^{(p-1)},\mu^{(p)},\lambda',\cdots,\lambda^{(n-1)}$ 不全为零. 这里 $\mu^{(p-1)}\neq 0$(因为 $\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-1)}$ 是线性无关的),实际上,$\lambda',\cdots,\lambda^{(n-1)}$ 不全为零(因为根据 1.4.1 节之(3),$\boldsymbol{\eta}^{(p-1)},\boldsymbol{\eta}^{(p)}$ 是线性无关的,即有 $\mu^{(p-1)}\boldsymbol{\eta}^{(p-1)}+\mu^{(p)}\boldsymbol{\eta}^{(p)}\neq\mathbf{0}$,导致 $\lambda'\boldsymbol{\xi}'+\cdots+\lambda^{(n-1)}\boldsymbol{\xi}^{(n-1)}\neq\mathbf{0}$). 因而,某个 λ(比如 $\lambda^{(n-1)}$)$\neq 0$,故 $\boldsymbol{\xi}^{(n-1)}=\boldsymbol{\eta}^{(p-1)},\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-2)}$ 的线性组合,并肯定包含 $\boldsymbol{\eta}^{(p-1)}$. 所以,(i) 可用 $\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-1)}$ 表达的所有矢量均可用 $\boldsymbol{\eta}^{(p-1)},\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-2)}$ 予以表达,以及(ii) $\boldsymbol{\eta}^{(p-1)},\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-2)}$ 是线性无关的(因为 $\boldsymbol{\eta}^{(p-1)}=\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-1)}$ 的线性组合,后者是线性无关的). 换句话说,$\boldsymbol{\eta}^{(p-1)},\boldsymbol{\eta}^{(p)},\boldsymbol{\xi}',\cdots,\boldsymbol{\xi}^{(n-2)}$ 形成 $\mathbf{V}_n$ 的一个基.

通过重复进行上述论证 p 次($p\leqslant n$),就可得上述定理的结果,即,$\boldsymbol{\eta}',\boldsymbol{\eta}'',\cdots,\boldsymbol{\eta}^{(p)}$($p\leqslant n$)可形成 $\mathbf{V}_n$ 的基的一部分. 证毕.

若在上述定理中令 $p=n$,即得下述引理.

引理　任何的 n 个线性无关矢量集合形成 $\mathbf{V}_n$ 的一个基.

而且,我们还已经证明了以下几个定理.

(3) 矢量空间和子空间的维数

定理　生成矢量空间 $\mathbf{V}_n$ 的线性无关矢量的个数总是 n,不依赖于基矢量的选择.

我们将这个数 n 称为此**矢量空间 $\mathbf{V}_n$ 的维数**.

对于子空间的维数,则有下述定理.

定理　对于由 $\mathbf{V}_n$ 中的矢量 $\boldsymbol{x}',\cdots,\boldsymbol{x}^{(m)}$ 所生成的矢量子空间 $\mathbf{S}$,生成此子空间 $\mathbf{S}$ 的线性无关矢量的个数是矢量集合 $\boldsymbol{x}',\cdots,\boldsymbol{x}^{(m)}$ 的线性无关矢量的个数(比如说)r. 这是**子空间 $\mathbf{S}$ 的维数**.

证　令 $\boldsymbol{x}',\cdots,\boldsymbol{x}^{(m)}$ 之中的线性无关矢量为(比如说)$\boldsymbol{x}',\cdots,\boldsymbol{x}^{(r)}$. 这些矢量形成 $\mathbf{S}$ 的一个基,因为 $\mathbf{S}$ 本身是一矢量空间,可将上面(2)中用 $\mathbf{S}$ 代替 $\mathbf{V}_n$ 予以重做,所以 $\mathbf{S}$ 的任何其他基也包含 $\mathbf{S}$ 的正好 r 个线性无关矢量.

尽管基的选择是任意的,但几何性质(例如,维数)是不变的!

1.5 线性变换与矢量空间映射的定理的明晰化

现在可将 1.3.3 节的定理进一步明晰化为下述形式.

定理 线性变换 A 将矢量空间 $\mathbf{V}_n$ 映射到由矢量 $\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)}$ 所生成的矢量子空间 $\mathbf{S}$,它是一个 r 维矢量空间,$r(\leqslant n)$为矢量集合 $\boldsymbol{a}',\cdots,\boldsymbol{a}^{(n)}$ 中线性无关矢量的个数.

现在我们能够按照 $r=n$ 或 $r<n$ 来区别两种类型的线性变换,同时进一步讨论两类线性变换的不同性质,并阐述与此有关的定理.

1.6 非奇异与奇异线性变换及有关定理

1.6.1 非奇异与奇异线性变换

定义 对于线性变换 $A=(\boldsymbol{a}',\cdots,\boldsymbol{a}^{(n)})$,若 $\boldsymbol{a}',\cdots,\boldsymbol{a}^{(n)}$ 是线性无关的($r=n$),则称该变换为**非奇异变换**;否则,若 $\boldsymbol{a}',\cdots,\boldsymbol{a}^{(n)}$ 是线性相关的($r<n$),则称该变换为**奇异变换**.

用这个定义和 1.5 节定理,我们看出下述结果.

1.6.2 线性变换的映射性质

一个非奇异线性变换将整个矢量空间 $\mathbf{V}_n$ 映射为其本身;而一个奇异变换则将 $\mathbf{V}_n$ 映射为 $\mathbf{V}_n$ 的真子空间 $\mathbf{S}$,$\mathbf{S}$ 具有维数 $r<n$. 或写成

非奇异变换: $\mathbf{V}_n\to\mathbf{V}_n$,

奇异变换: $\mathbf{V}_n\to\mathbf{S}(r<n)$.

非奇异线性变换与奇异线性变换之间的这个区别是相当重要的(量变引起质变),因为非奇异线性变换具有一个特殊性质,它是奇异线性变换所不能分享的. 这里所指的是非奇异线性变换存在逆变换,如下面两个定理所示.

1.6.3 非奇异线性变换的一一映射性质

定理 非奇异线性变换 A 所作映射

$$\boldsymbol{x}\to\boldsymbol{y}=A\boldsymbol{x}$$

是**一一映射**.

证 这里只需要证明,对于一个给定的 $\boldsymbol{y}$,有一个且仅有一个 $\boldsymbol{x}$ 与之对应. 既然 $\boldsymbol{a}',\cdots,\boldsymbol{a}^{(n)}$ 是线性无关的,所以按照 1.4.2 节(2)中引理,它们形成 $\mathbf{V}_n$ 中的一个

基,从而,给定的 $\boldsymbol{y}$ 可用它们加以表达.事实上仅有一种方式来表达 $\boldsymbol{y}=\mu'\boldsymbol{a}'+\cdots+\mu^{(n)}\boldsymbol{a}^{(n)}$.因为如果有两种方式来表达,比如说,

$$\boldsymbol{y}=\mu'\boldsymbol{a}'+\cdots+\mu^{(n)}\boldsymbol{a}^{(n)}=\lambda'\boldsymbol{a}'+\cdots+\lambda^{(n)}\boldsymbol{a}^{(n)},$$

则两者相减得

$$(\mu'-\lambda')\boldsymbol{a}'+\cdots+(\mu^{(n)}-\lambda^{(n)})\boldsymbol{a}^{(n)}=\boldsymbol{0},$$

而根据矢量 $\boldsymbol{a}',\cdots,\boldsymbol{a}^{(n)}$ 线性无关的定义(1.4.1节)必然有

$$\mu'-\lambda'=\cdots=\mu^{(n)}-\lambda^{(n)}=0 \quad 即 \quad \mu'=\lambda',\cdots,\mu^{(n)}=\lambda^{(n)},$$

所以 $\mu',\cdots,\mu^{(n)}$ 是唯一的.

又因

$$\boldsymbol{a}'=A\boldsymbol{\xi}',\quad \cdots,\quad \boldsymbol{a}^{(n)}=A\boldsymbol{\xi}^{(n)},$$

直接计算容易证明

$$\begin{aligned}\boldsymbol{y}&=\mu'\boldsymbol{a}'+\cdots+\mu^{(n)}\boldsymbol{a}^{(n)}\\&=\mu'A\boldsymbol{\xi}'+\cdots+\mu^{(n)}A\boldsymbol{\xi}^{(n)}\\&=A\mu'\boldsymbol{\xi}'+\cdots+A\mu^{(n)}\boldsymbol{\xi}^{(n)}\\&=A(\mu'\boldsymbol{\xi}'+\cdots+\mu^{(n)}\boldsymbol{\xi}^{(n)}),\end{aligned}$$

它可读作 $A\boldsymbol{x}$,即 $\boldsymbol{y}=A\boldsymbol{x}$,具有

$$\boldsymbol{x}=\mu'\boldsymbol{\xi}'+\cdots+\mu^{(n)}\boldsymbol{\xi}^{(n)},$$

它是唯一的.证毕.

1.6.4 非奇异线性变换具有逆变换

定理 对于非奇异线性变换 $\boldsymbol{x}\to\boldsymbol{y}=A\boldsymbol{x}$,其逆映射 $\boldsymbol{y}\to\boldsymbol{x}$ 本身也是线性变换.

证 由于,对 $\boldsymbol{y}$ 的特殊选择,$\boldsymbol{y}=\begin{bmatrix}1\\0\\0\\\vdots\\0\end{bmatrix}=\boldsymbol{\xi}'$,按照1.6.3节定理,存在一个 $\boldsymbol{x}$ 与之对应,比如说,$\boldsymbol{x}=\boldsymbol{b}'$,即 $A\boldsymbol{b}'=\boldsymbol{\xi}'$.同样,对 $\boldsymbol{y}$ 的特殊选择,$\boldsymbol{y}=\begin{bmatrix}0\\1\\0\\\vdots\\0\end{bmatrix}=\boldsymbol{\xi}''$,按照1.6.3节定理,存在一个 $\boldsymbol{x}$ 与之对应,比如说,$\boldsymbol{x}=\boldsymbol{b}''$,即 $A\boldsymbol{b}''=\boldsymbol{\xi}''$.如此等等,直至,

根据同样理由，对 $\boldsymbol{y}$ 的特殊选择，$\boldsymbol{y}=\begin{bmatrix}0\\0\\\vdots\\0\\1\end{bmatrix}=\boldsymbol{\xi}^{(n)}$，存在一个 $\boldsymbol{x}$ 与之对应，比如说，$\boldsymbol{x}=\boldsymbol{b}^{(n)}$，即 $A\boldsymbol{b}^{(n)}=\boldsymbol{\xi}^{(n)}$. 这样求得的 $\boldsymbol{b}',\boldsymbol{b}'',\cdots,\boldsymbol{b}^{(n)}$ 满足 $A\boldsymbol{b}^{(i)}=\boldsymbol{\xi}^{(i)}(i=1,2,\cdots,n)$，按照 1.6.3 节定理，它们是唯一的. 因此，对任意的 $y_1,y_2,\cdots,y_n$，有

$$\begin{aligned}A(y_1\boldsymbol{b}'+y_2\boldsymbol{b}''+\cdots+y_n\boldsymbol{b}^{(n)})&=y_1A\boldsymbol{b}'+y_2A\boldsymbol{b}''+\cdots+y_nA\boldsymbol{b}^{(n)}\\&=y_1\boldsymbol{\xi}'+y_2\boldsymbol{\xi}''+\cdots+y_n\boldsymbol{\xi}^{(n)}\\&=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}=\boldsymbol{y},\end{aligned}$$

即求得与 $\boldsymbol{y}$ 对应的 $\boldsymbol{x}$ 为

$$\boldsymbol{x}=y_1\boldsymbol{b}'+y_2\boldsymbol{b}''+\cdots+y_n\boldsymbol{b}^{(n)}.$$

因而，对于非奇异线性变换 A，映射

$$\boldsymbol{x}\to\boldsymbol{y}=A\boldsymbol{x}=x_1\boldsymbol{a}'+x_2\boldsymbol{a}''+\cdots+x_n\boldsymbol{a}^{(n)},$$

存在唯一的逆映射

$$\boldsymbol{y}\to\boldsymbol{x}=B\boldsymbol{y}=y_1\boldsymbol{b}'+y_2\boldsymbol{b}''+\cdots+y_n\boldsymbol{b}^{(n)},$$

它也是一个线性变换. 这里

$$\begin{aligned}A&=(\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)})=\{a_i^{(k)}\}=\{a_{ik}\},\\B&=(\boldsymbol{b}',\boldsymbol{b}'',\cdots,\boldsymbol{b}^{(n)})=\{b_i^{(k)}\}=\{b_{ik}\}.\end{aligned}$$

于是映射 $\boldsymbol{x}\to\boldsymbol{y}$ 还可写出为

$$y_i=\sum_{k=1}^{n}a_{ik}x_k\quad(i=1,2,\cdots,n),$$

而逆映射 $\boldsymbol{y}\to\boldsymbol{x}$ 则可写出为

$$x_i=\sum_{k=1}^{n}b_{ik}y_k\quad(i=1,2,\cdots,n).$$

通常将 B 用 A^{-1} 来表示，并将 $\boldsymbol{y}\to\boldsymbol{x}=A^{-1}\boldsymbol{y}$ 称为 $\boldsymbol{x}\to\boldsymbol{y}=A\boldsymbol{x}$ 之**逆变换**.

还应注意到，对于非奇异线性变换 A 有一一映射

$$\boldsymbol{0}\leftrightarrow\boldsymbol{0}$$

$$\boldsymbol{y}=A\boldsymbol{x}\leftrightarrow\boldsymbol{x}=A^{-1}\boldsymbol{y}.$$

1.6.5 奇异线性变换的情况

对于奇异线性变换

$$x \to y = Ax = x_1 a' + x_2 a'' + \cdots + x_n a^{(n)},$$

情况完全不同.

因为非奇异线性变换 $A=[a',a'',\cdots,a^{(n)}]$,矢量 $a',a'',\cdots,a^{(n)}$ 是线性无关的,有唯一的逆变换,它们是一一映射,就是说,总有

$$x \leftrightarrow y \quad 导致 \quad \mathbf{V}_n \leftrightarrow \mathbf{V}_n.$$

而对于奇异线性变换 $A=[a',a'',\cdots,a^{(n)}]$,矢量 $a',a'',\cdots,a^{(n)}$ 是线性相关的,即 $\lambda_1 a'+\lambda_2 a''+\cdots+\lambda_n a^{(n)}$ 会在 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 不全为零的情况下变为零;如果矢量 $a',\cdots,a^{(n)}$ 中线性无关矢量的个数为 $r(r<n)$,则

$$x \to y \quad 导致 \quad \mathbf{V}_n \to \mathbf{S},$$

这里 $\mathbf{S}$ 是 n 维矢量空间 $\mathbf{V}_n$ 中的 r 维矢量子空间. 所以,此变换是多到一映射.

至于奇异线性变换的逆变换问题,如果 y 在 $\mathbf{S}$ 中(或者说 $y\in\mathbf{S}$),则有逆变换但不唯一,

$$y \to x \quad 导致 \quad \mathbf{S} \to \mathbf{V}_n,$$

这是一到多映射;但是,如果 y 不在 $\mathbf{S}$ 中(或者说 $y\in \mathbf{V}_n-\mathbf{S}$),则不存在逆变换,就是说,$\mathbf{V}_n-\mathbf{S}\to\mathbf{V}_n$ 是不可能的. 应该注意到,$\mathbf{S}$ 中总是包含有零矢量 $\mathbf{0}$. 因而总会有 $\mathbf{0}\leftrightarrow\mathbf{0}$.

例如:具体地说,若

$$A = [a',a'',a'''] = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 2 & 4 \end{bmatrix},$$

显然有 $a'''=2a''$,所以 A 是奇异变换;又由于 $\lambda' a'+\lambda'' a''=\mathbf{0}$ 导致

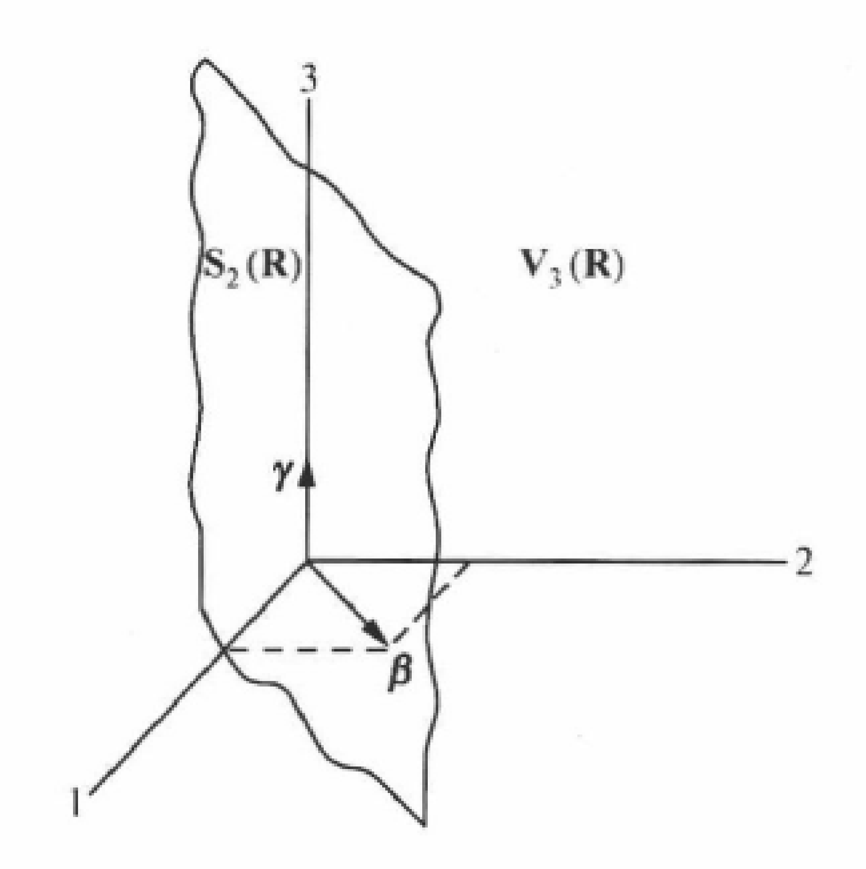
$$\left.\begin{aligned} \lambda' + \lambda'' &= 0, \\ \lambda' + 2\lambda'' &= 0 \end{aligned}\right\} \Longrightarrow \lambda' = \lambda'' = 0,$$

所以 a',a''是线性无关的,即 $r=2$. 因此(为了方便,下面将 $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ 记作$[a,b,c]^{\mathrm{T}}$),

$$x = [x_1,x_2,x_3]^{\mathrm{T}} \to y = [b,b,c]^{\mathrm{T}} \quad 即 \quad \mathbf{V}_3(\mathbf{R}) \to \mathbf{S}_2(\mathbf{R}),$$

这里 $\mathbf{S}_2(\mathbf{R})=[\boldsymbol{\beta},\boldsymbol{\gamma}]$是 $\mathbf{V}_3(\mathbf{R})$中由 $\boldsymbol{\beta}=[1,1,0]^{\mathrm{T}}$ 和 $\boldsymbol{\gamma}=[0,0,1]^{\mathrm{T}}$ 所生成的子空间;如下页图所示.

现在考虑逆变换问题. 若给定 $y=[b,b,c]^{\mathrm{T}}\in\mathbf{S}_2$,则有 $x_1=2b-c$ 和 $x_2+2x_3=c-b$,这就是说,$\mathbf{S}_2$ 中一点(一个矢量)对应于 $\mathbf{V}_3$ 中两平面相交的一条直线(无穷个



矢量),是一到多(无穷)映射.[相反,若给定 $\boldsymbol{x}$ 位于此直线上任一点,显然会给出 $\mathbf{S}_2$ 中一点 $\boldsymbol{y}=[b,b,c]^{\mathrm{T}}$,是多(无穷)到一映射.]另一方面,若给定 $\boldsymbol{y}$ 的 $y_1 \neq y_2$($\boldsymbol{y} \in \mathbf{V}_3 - \mathbf{S}_2$),则导致方程矛盾而无解.

第 2 章　群

2.1　非奇异线性变换总体的性质

在这一章中，我们将要考虑线性空间 $\mathbf{V}_n$ 的全体非奇异线性变换 A, B, C，等等：

$$A = [\boldsymbol{a}', \boldsymbol{a}'', \cdots, \boldsymbol{a}^{(n)}],$$
$$B = [\boldsymbol{b}', \boldsymbol{b}'', \cdots, \boldsymbol{b}^{(n)}],$$
$$C = [\boldsymbol{c}', \boldsymbol{c}'', \cdots, \boldsymbol{c}^{(n)}],$$

等等；这里 $\boldsymbol{a}', \boldsymbol{a}'', \cdots, \boldsymbol{a}^{(n)}$ 是线性无关的，$\boldsymbol{b}', \boldsymbol{b}'', \cdots, \boldsymbol{b}^{(n)}$ 是线性无关的，等等.

2.1.1　非奇异线性变换具有逆变换

从上一章 1.6.1～1.6.4 各节可以总结出，对于每个非奇异线性变换，比如说 A，可建立一个一一映射：

$$\boldsymbol{x} \to \boldsymbol{y} = A\boldsymbol{x},$$

它将矢量空间 $\mathbf{V}_n$ 中的任一 $\boldsymbol{x}$ 映射为子空间 $\mathbf{S}$ 中的唯一 $\boldsymbol{y}$；同时存在**逆变换** A^{-1}，会将 $\mathbf{S}$ 中的该 $\boldsymbol{y}$ 逆映射回原 $\boldsymbol{x}$，也是一一映射：

$$\boldsymbol{y} \to \boldsymbol{x} = A^{-1}\boldsymbol{y};$$

并且 $\mathbf{S}$ 与 $\mathbf{V}_n$ 重合，或者说，非奇异线性变换

$$\boldsymbol{x} \leftrightarrow \boldsymbol{y} \quad 导致 \quad \mathbf{V}_n \leftrightarrow \mathbf{V}_n.$$

2.1.2　非奇异线性变换具有恒同变换

定义　我们将由单位矩阵

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

即对角元均为 1 而非对角元均为 0 的方阵，所描述的变换

$$\boldsymbol{x} \to I\boldsymbol{x} = \boldsymbol{x},$$

称为**恒同线性变换**.

2.1.3 线性变换之积

定义 对于由矩阵 A 和 B 所描述的两个非奇异线性变换，由矩阵积 BA 所描述的线性变换称为该两**变换之积**.

因为对于

$$\boldsymbol{x} \to \boldsymbol{y} = A\boldsymbol{x},$$
$$\boldsymbol{y} \to \boldsymbol{z} = B\boldsymbol{y},$$

通过直接代换可以得到

$$\boldsymbol{x} \to \boldsymbol{z} = B(A\boldsymbol{x}) = (BA)\boldsymbol{x},$$

这里应用了矩阵运算之结合律；由此可见，矩阵积 BA 的确描述了

$$\boldsymbol{x} \xrightarrow{A} \boldsymbol{y} \xrightarrow{B} \boldsymbol{z}$$

之积变换

$$\boldsymbol{x} \xrightarrow{BA} \boldsymbol{z}.$$

首先应该注意到，由于矩阵不满足交换律，$BA \neq AB$，所以，这个定义依赖于施行变换的先后次序；因为 A 先 B 后时有

$$\boldsymbol{x} \xrightarrow{A} \boldsymbol{y} \xrightarrow{B} \boldsymbol{z} \text{ 之积变换为 } \boldsymbol{x} \xrightarrow{BA} \boldsymbol{z},$$

而 B 先 A 后时则有

$$\boldsymbol{x} \xrightarrow{B} \boldsymbol{y}' \xrightarrow{A} \boldsymbol{z}' \text{ 之积变换为 } \boldsymbol{x} \xrightarrow{AB} \boldsymbol{z}';$$

可见其结果一般情况下是不同的.

另外，因为 A 和 B 都是描述非奇异线性变换，有一一映射性质

$$A \text{ 和 } A^{-1}: \quad \boldsymbol{x} \leftrightarrow \boldsymbol{y} \quad \text{导致} \quad \mathbf{V}_n \leftrightarrow \mathbf{V}_n,$$
$$B \text{ 和 } B^{-1}: \quad \boldsymbol{y} \leftrightarrow \boldsymbol{z} \quad \text{导致} \quad \mathbf{V}_n \leftrightarrow \mathbf{V}_n,$$

所以，$C = BA$ 有

$$\boldsymbol{x} \xrightarrow{A} \boldsymbol{y} \xrightarrow{B} \boldsymbol{z} \quad \text{导致} \quad \mathbf{V}_n \to \mathbf{V}_n,$$

而其逆变换 C^{-1} 应为

$$\boldsymbol{x} \xleftarrow{A^{-1}} \boldsymbol{y} \xleftarrow{B^{-1}} \boldsymbol{z} \quad \text{导致} \quad \mathbf{V}_n \leftarrow \mathbf{V}_n.$$

即 $C = BA$ 有 $C^{-1} = (BA)^{-1} = A^{-1}B^{-1}$ 也具有一一映射性质

$$C \text{ 和 } C^{-1}: \boldsymbol{x} \leftrightarrow \boldsymbol{z} \quad \text{导致} \quad \mathbf{V}_n \leftrightarrow \mathbf{V}_n,$$

也就是说，$C = BA$ 也是描述非奇异线性变换. 注意到乘积之逆为因子之逆沿相反次序之积，这从

$$\boldsymbol{z} = B\boldsymbol{y}, \boldsymbol{y} = A\boldsymbol{x} \quad \text{得} \quad \boldsymbol{z} = (BA)\boldsymbol{x} = C\boldsymbol{x},$$

和

$$\boldsymbol{x} = A^{-1}\boldsymbol{y}, \boldsymbol{y} = B^{-1}\boldsymbol{z} \quad 得 \quad \boldsymbol{x} = (A^{-1}B^{-1})\boldsymbol{z} = C^{-1}\boldsymbol{z}$$

可见.

2.1.4 线性变换的乘法满足结合律

组成线性变换之积的这种方式(称为**乘法**)也满足结合律. 因为,若 A,B,C 均为线性空间 $\mathbf{V}_n$ 中描述非奇异线性变换的矩阵,

$$A: \boldsymbol{x} \to \boldsymbol{y} = A\boldsymbol{x},$$
$$B: \boldsymbol{y} \to \boldsymbol{z} = B\boldsymbol{y},$$
$$C: \boldsymbol{z} \to \boldsymbol{t} = C\boldsymbol{z};$$

则按照线性变换之积的定义显然有

$$(BA): \boldsymbol{x} \to \boldsymbol{z} = (BA)\boldsymbol{x},$$
$$C(BA): \boldsymbol{x} \to \boldsymbol{t} = C(BA)\boldsymbol{x};$$

另一方面又有

$$(CB): \boldsymbol{y} \to \boldsymbol{t} = (CB)\boldsymbol{y},$$
$$(CB)A: \boldsymbol{x} \to \boldsymbol{t} = (CB)A\boldsymbol{x};$$

而按照矩阵乘法的结合律(1.2.2 节之(5))$C(BA)=(CB)A$,即得**线性变换的乘法满足结合律**.

2.1.5 非奇异线性变换的几何意义

利用“转动”的类比来看变换的**几何意义**.

恒同变换 I 是将矢量 $\boldsymbol{x}$ 放在原处不动(见下图(i)):

$$\boldsymbol{x} \to \boldsymbol{x} = I\boldsymbol{x}.$$

非奇异线性变换 A 是将矢量 $\boldsymbol{x}$“转”一“角度”而变到矢量 $\boldsymbol{y}$ 的位置(见下图(ii))

$$\boldsymbol{x} \to \boldsymbol{y} = A\boldsymbol{x}.$$

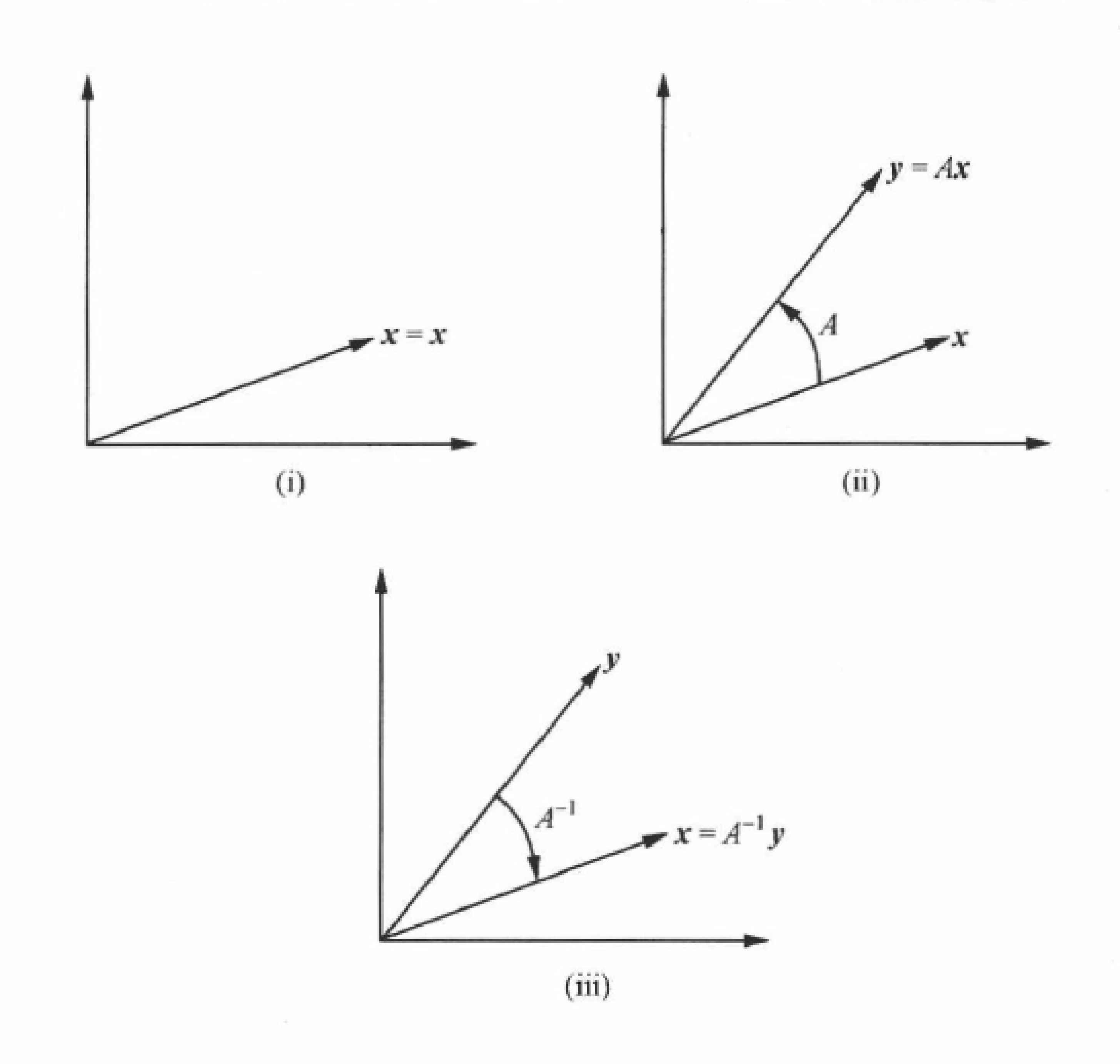
非奇异线性变换 A 的逆变换 A^{-1} 是将矢量 $\boldsymbol{y}$“转”回到原矢量 $\boldsymbol{x}$ 的位置(见下图(iii)):

$$\boldsymbol{y} \to \boldsymbol{x} = A^{-1}\boldsymbol{y}.$$

所以,$A^{-1}A$ 是表示先将 $\boldsymbol{x}$“转”到 $\boldsymbol{y}$,再将 $\boldsymbol{y}$“转”回到 $\boldsymbol{x}$,结果等于 $\boldsymbol{x}$ 不动,即

$$A^{-1}A = I.$$

为了进一步深入讨论,下面引进抽象群的重要概念.



(i) (ii) (iii)

2.2 抽象群的定义[①②]

2.2.1 定义

设有一个由元素 $a,b,c,\cdots$ 组成的集合 G，G 中元素对于一个单值二元运算（用 $\circ$ 表示，称为乘法）满足下列条件：

(i) 封闭性.

若 a,b 是 G 中任意元素，则有唯一确定的 $a\circ b=c$，而 c 也是 G 中元素.

(ii) 满足结合律.

若 a,b,c 是 G 中任意元素，则有

$$(a\circ b)\circ c=a\circ(b\circ c).$$

(iii) 有单位元素存在.

G 中有一单位元素 1［有时用 e（德文 Einheit）或 I（英文 Identity）表示］，使得对 G 中任意元素 a 都有

$$1\circ a=a=a\circ 1.$$

---

① G. Birkhoff, S. MacLane, *A Survey of Modern Algebra*, pp. 130—131.

② 参考：《中国大百科全书：数学》. 546 页，群（段学复撰），中国大百科全书出版社，北京，1988.

(iv) 有逆元素存在.

对于 G 中每一元素 a 有一元素 a^{-1} 使得

$$a^{-1} \circ a = 1 = a \circ a^{-1},$$

a^{-1} 称为元素 a 的逆元素.

这样的集合 G 对该乘法运算 $\circ$ 称为**群**.

若群 G 对乘法还满足交换律,即对 G 中任意两个元素 a,b 都有 $a \circ b = b \circ a$,则 G 称为**交换群**或 **Abel 群**.

附注一 其实,上述条件中列入 $a \circ 1 = a$ 和 $a \circ a^{-1} = 1$ 并非必要,因为可以证明,它们是其他几个条件的必然结果.

给定 $(a \circ b) \circ c = a \circ (b \circ c)$,$1 \circ a = a$ 和 $a^{-1} \circ a = 1$,则蕴涵① $a \circ 1 = a$ 和② $a \circ a^{-1} = 1$.

证 ① $\because$ $a^{-1} \circ (a \circ 1) = (a^{-1} \circ a) \circ 1 = 1 \circ 1 = 1 = a^{-1} \circ a$,所以再以 $(a^{-1})^{-1}$ 左乘上式两端得到

$$1 \circ (a \circ 1) = 1 \circ a, \quad \therefore a \circ 1 = a.$$

② $\because$ $a^{-1} \circ (a \circ a^{-1}) = (a^{-1} \circ a) \circ a^{-1} = 1 \circ a^{-1} = a^{-1}$,所以再以 $(a^{-1})^{-1}$ 左乘上式两端得到

$$1 \circ (a \circ a^{-1}) = 1, \quad \therefore a \circ a^{-1} = 1.$$

附注二 应该注意到,对于 G 中任意二元素之积 $a \circ b$,其逆元素为

$$(a \circ b)^{-1} = b^{-1} \circ a^{-1};$$

或者更一般地有

$$(a \circ b \circ c \circ \cdots \circ f)^{-1} = f^{-1} \circ \cdots \circ c^{-1} \circ b^{-1} \circ a^{-1}.$$

证 因为前式(上式的前一式)两边同时右乘以 $(a \circ b)$ 后左边为 1,而右边为

$$\begin{aligned}(b^{-1} \circ a^{-1}) \circ (a \circ b) &= b^{-1} \circ (a^{-1} \circ a) \circ b = b^{-1} \circ 1 \circ b \\ &= b^{-1} \circ b = 1,\end{aligned}$$

所以前式两边相等,得证. 对于一般结果,可以类似地加以证明.

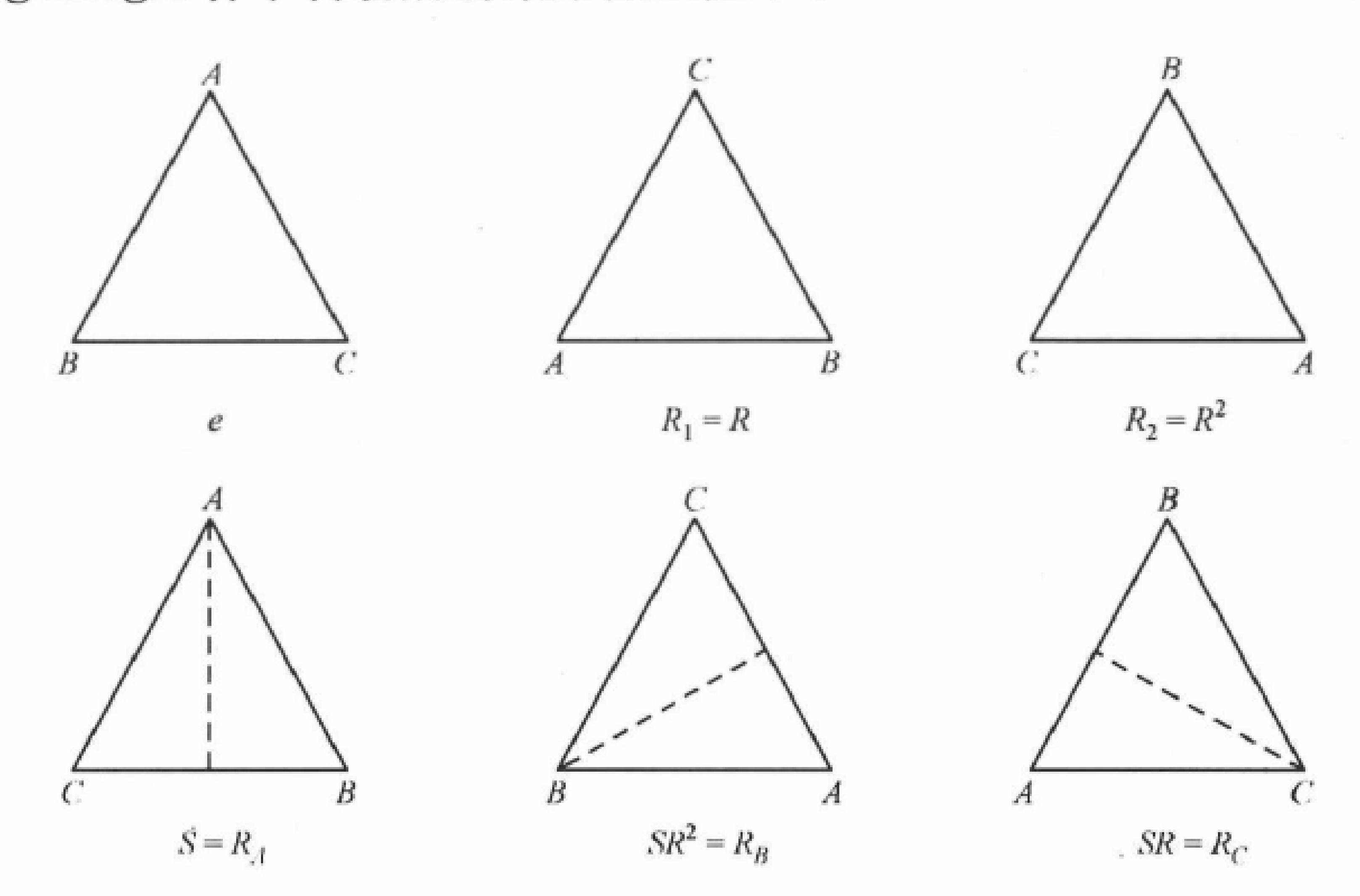
2.2.2 说明与例子

说明:群是数学中最基本的集合,同时从上述的群的性质可以推演出很多漂亮的结果来,所以,群是非常重要的概念.

群又往往和"**对称性**"发生联系. 凡是具有对称性的东西,施加某种运算后可使其外表保持不变,对于同一种对称性的许多种运算即满足群的定义,所以形成一个群.

例如,一个等边三角形具有对称心(中心)和对称轴(三顶点至对边的垂线),保持其空间位置不变的刚体运动共有下列 6 种:绕对称心转动 $\frac{2k\pi}{3}$ $(k=0,1,2)$ 的 3

个转动 e,R,R^2[转动 0π 为不动，转动 $l(2\pi)$(l 为整数)复原与不动状态相同，故 $R^3=R^0=e$]；和绕三个对称轴分别转动 π 的 3 个转动 S,SR,SR^2(相当于反射，S 来自德文 Spiegelung). 各个转动后的结果图示如下：



这 6 种运动对于相继实施的运算构成一个群，称为 3 边的**二面体群 $\mathbf{D}_3$**.

为了证明这一点，我们来建立此 6 个元素的乘法表，见表 2.1.

表 2.1 $\mathbf{D}_3$ 群乘法表

a \ b ($a\circ b$)	e	R	R^2	S	SR	SR^2
e	e	R	R^2	S	SR	SR^2
R^2	R^2	e	R	SR	SR^2	S
R	R	R^2	e	SR^2	S	SR
S	S	SR	SR^2	e	R	R^2
SR	SR	SR^2	S	R^2	e	R
SR^2	SR^2	S	SR	R	R^2	e

表 2.1 是这样建立的：先实施第一行中某一运算，再实施第一列中某一运算，则在对应行和列的相交处就是其乘积. 注意到下列基本关系：

$$R^3=e,\quad S^2=e,\quad (SR^k)(SR^k)=e,\quad SR^kS=R^{3-k};$$

不难进行验证.(此结果可推广至 n 边的二面体群 $\mathbf{D}_n$，只要令 $3\to n, k=0,\cdots,n-1$ 即可.)

有了这个乘法表，就很容易证明

$$e,\quad R,\quad R^2,\quad S,\quad SR,\quad SR^2$$

这 6 个元素构成一个群. 同时，e,R,R^2 这 3 个元素也构成一个群. 因为可以看出它

们的确满足：

(i) 封闭性. 明显易见.

(ii) 结合律. 这是线性变换，满足结合律；利用上述基本关系，也可直接进行验证.

(iii) 有单位元素 e 存在. 明显易见.

(iv) 有逆元素存在. 很明显，相交点是 e 的两元素互为逆元素.

2.2A 附录：域的另一定义

若一集合 **F** 中的元素，对加法和(除 0 元素外)对乘法都构成交换群，且乘法对加法有分配律，则这样的集合就称为**域**.

域的这个定义与第 1 章 1.1A 节附录中的定义是等价的. 因为按照群的条件(i)封闭性，**F** 中任意二元素之和与积亦含于 **F** 中；(ii) 满足结合律，对加法是 $a+(b+c)=(a+b)+c$，对乘法是$a\cdot(b\cdot c)=(a\cdot b)\cdot c$，即 $a(bc)=(ab)c$；(iii) 有单位元素存在，对加法是 0 元素 $0+a=a$，对乘法是 1 元素 $1a=a$；(iv) 有逆元素存在，对加法 a 的逆元素是 $-a$，$(-a)+a=0$，对乘法(除 0 外)a 的逆元素是 a^{-1}，$a^{-1}a=1$；另外，交换群还要求满足交换律，$a+b=b+a$ 及 $ab=ba$，而且上面关于域的新定义中还要求乘法对加法有分配律 $a(b+c)=ab+ac$. 所以，两个定义是完全等价的.

下面将要讨论各种各样具体的群，以及与之相关的一些概念.

2.3 一般线性群

2.3.1 线性变换群

首先，我们将 2.1 节关于非奇异线性变换总体的性质总结一下并与群的定义进行对比如下：

(i) 两个非奇异线性变换之积仍是一个非奇异线性变换，满足群元素的封闭性条件.

(ii) 对于非奇异线性变换之乘积，结合律成立，满足对群元素乘积的同样要求.

(iii) 存在一个恒同非奇异线性变换 1，当它与任意一个非奇异线性变换 A 相乘时仍然得到 A，即 $1A=A1=A$，也就是说，1 是群的定义中的单位元素.

(iv) 对于任意一个非奇异线性变换 A，存在一个非奇异线性逆变换 A^{-1}，使得 $A^{-1}A=AA^{-1}=1$，满足群元素中存在逆元素的条件.

因此,非奇异线性变换总体构成一个群,称为**线性变换群**.

2.3.2 矩阵群

全体非奇异 n 阶矩阵相对于链式乘法构成一个群(所谓非奇异 n 阶矩阵 $A=[\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)}]$,意指其列矢量 $\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)}$ 是线性无关的),称为**矩阵群**.

这是因为描述非奇异线性变换的矩阵是非奇异矩阵,相应地具有下列性质:

(i) 任意两个非奇异 n 阶矩阵 A,B 之积 AB 仍是非奇异 n 阶矩阵(见 2.1.3 节).

(ii) 任意三个非奇异 n 阶矩阵 A,B,C 之积满足结合律:$(AB)C=A(BC)=ABC$.

注意:这里假设 n 为有限.如果 n 为无穷大,则上式包含两个极限的次序变换;而结合律是否成立,依赖于这个次序变换是否成立.这个问题需要另行考虑.

(iii) 有 n 阶单位矩阵

$$I=\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}$$

存在(见 2.1.2 节),因为对任意非奇异 n 阶矩阵 A 显然有

$$IA=AI=A.$$

(iv) 对任意非奇异 n 阶矩阵 A 有逆矩阵 A^{-1},它也是非奇异 n 阶矩阵(见 2.1.1 节),而显然有

$$A^{-1}A=AA^{-1}=I.$$

这些性质正好是构成群所应满足的条件.

2.3.3 群的同构

定义 如果群 G 的元素与群 G^* 的元素是一一对应的,即有

$$a\leftrightarrow a^*,$$
$$b\leftrightarrow b^*,$$
$$ab\leftrightarrow a^*b^*;$$

则称群 G 与群 G^* 为**同构**,记作

$$G\cong G^*.$$

容易看出,这个定义必然蕴涵下面关于单位元素与逆元素的对应关系

$$e\leftrightarrow e^*,$$
$$a^{-1}\leftrightarrow(a^*)^{-1};$$

只要令定义中的 b 为 e 或 a^{-1} 即可得出.

应该注意,群的同构概念中强调的是其元素以及其积元素之间的一一对应关系,并未涉及具体元素或具体运算,后者可以完全不同.

2.3.4　一般线性群

因为非奇异线性变换与非奇异矩阵是一一对应的,所以 $\mathbf{V}_n(\mathbf{F})$ 中的非奇异线性变换(其积由矩阵积予以定义)所构成的群(线性变换群)与域 $\mathbf{F}$ 上的 n 阶非奇异矩阵(相对于链式乘法)所构成的群(矩阵群)是"同构的",通称为域 $\mathbf{F}$ 上的**一般线性群**,记作 $\mathbf{GL}_n(\mathbf{F})$. $\mathbf{GL}_n(\mathbf{F})$ 中所有行列式(关于行列式见第 3 章)为 $+1$ 的矩阵对于乘法也构成群,称为域 $\mathbf{F}$ 上的**特殊线性群**,记作 $\mathbf{SL}_n(\mathbf{F})$.

2.3.5　连续群

一般线性群 $\mathbf{GL}_n(\mathbf{F})$ 中的元素数是无限的,它是 $n\times n$ 个参数 a_{ik} 的**连续群**,每个 a_{ik} 都在域 $\mathbf{F}$ 上变化.(当然,需满足非奇异性的要求.)

2.4　仿射变换群

2.4.1　子群①

有许多群被包含于较大的群之中,如我们将要在以下几节中予以阐明的. 现在先引进子群的概念.

定义　若群 G 的元素集合中的一个非空子集合 H,它对于群 G 的运算自成一群,则 H 称为 G 的一个**子群**.

例如,2.2.2 节中所举二面体群的 6 个元素 e,R_1,R_2,R_A,R_B,R_C 中的前 3 个元素 e,R_1,R_2 对于绕对称心的转动自成一群.

2.4.2　仿射变换群

定义　对于数域 $\mathbf{F}$ 上的非奇异 n 阶方阵 $A=\{a_{ik}\}$ 和矢量空间 $\mathbf{V}_n(\mathbf{F})$ 中的常矢量 $\boldsymbol{l}$ 为给定的情况下,将 $\mathbf{V}_n(\mathbf{F})$ 中的矢量 $\boldsymbol{x}$ 变换为矢量 $\boldsymbol{y}$ 的下列变换:

$$\boldsymbol{x}\rightarrow\boldsymbol{y}=A\boldsymbol{x}+\boldsymbol{l},$$

或者

$$y_i=\sum_{k=1}^{n}a_{ik}x_k+l_i\quad(i=1,2,\cdots,n),$$

① G. Birkhoff, S. MacLane, *A Survey of Modern Algebra*, p. 143.

展开写出即为

$$y_1 = a_{11}x_1 + \cdots + a_{1n}x_n + l_1,$$
$$\vdots$$
$$y_n = a_{n1}x_1 + \cdots + a_{nn}x_n + l_n,$$

称为 $\mathbf{V}_n(\mathbf{F})$上的一个**仿射变换**.

仿射变换最重要的性质是保持点的共线性(或共面性)以及保持直线的平行性.这里不作详细讨论.

$\mathbf{V}_n(\mathbf{F})$的全体仿射变换构成一个群,称为**仿射变换群**,记作 $\mathbf{Aff}_n(\mathbf{F})$.

证 若有下列任意三个仿射变换:

$$\boldsymbol{x} \to \boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{l},$$
$$\boldsymbol{y} \to \boldsymbol{z} = B\boldsymbol{y} + \boldsymbol{p},$$
$$\boldsymbol{z} \to \boldsymbol{t} = C\boldsymbol{z} + \boldsymbol{q}.$$

首先,它们相继进行的结果可以得到

$$\boldsymbol{x} \to \boldsymbol{z} = B(A\boldsymbol{x} + \boldsymbol{l}) + \boldsymbol{p} = (BA)\boldsymbol{x} + (B\boldsymbol{l} + \boldsymbol{p}),$$
$$\boldsymbol{y} \to \boldsymbol{t} = C(B\boldsymbol{y} + \boldsymbol{p}) + \boldsymbol{q} = (CB)\boldsymbol{y} + (C\boldsymbol{p} + \boldsymbol{q});$$

由于 BA 仍是非奇异 n 阶方阵,$B\boldsymbol{l} + \boldsymbol{p}$ 仍是一个常矢量,它仍是一个仿射变换,所以满足群元素的封闭性条件.

另一方面,

$$(\boldsymbol{x} \to \boldsymbol{z}) \to \boldsymbol{t}:\ \boldsymbol{t} = C\{(BA)\boldsymbol{x} + (B\boldsymbol{l} + \boldsymbol{p})\} + \boldsymbol{q}$$
$$= C(BA)\boldsymbol{x} + \{C(B\boldsymbol{l}) + C\boldsymbol{p} + \boldsymbol{q}\},$$
$$\boldsymbol{x} \to (\boldsymbol{y} \to \boldsymbol{t}):\boldsymbol{t} = \{(CB)\boldsymbol{y} + (C\boldsymbol{p} + \boldsymbol{q})\}$$
$$= (CB)\{A\boldsymbol{x} + \boldsymbol{l}\} + \{C\boldsymbol{p} + \boldsymbol{q}\}$$
$$= (CB)A\boldsymbol{x} + \{(CB)\boldsymbol{l} + C\boldsymbol{p} + \boldsymbol{q}\}.$$

由于矩阵满足结合律$(CB)A = C(BA)$,$(CB)\boldsymbol{l} = C(B\boldsymbol{l})$,所以仿射变换满足结合律.

存在恒同变换由单位矩阵 I 和常矢量 $\mathbf{0}$ 给出,$\boldsymbol{x} \to \boldsymbol{x} = I\boldsymbol{x} + \mathbf{0}$,这是群中的单位元素.

最后,对于任意仿射变换

$$\boldsymbol{x} \to \boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{l},$$

存在逆变换

$$\boldsymbol{y} \to \boldsymbol{x} = A^{-1}\boldsymbol{y} - A^{-1}\boldsymbol{l},$$

此即该仿射变换的逆元素.

所以,全体仿射变换满足群的全部条件.

2.4.3 仿射变换群的子群

仿射变换群包含两个子群:

(i) 一般线性群 $\mathbf{GL}_n(\mathbf{F})$(相当于常矢量 $\boldsymbol{l}$ 恒为 $\mathbf{0}$)：

$$\boldsymbol{x} \rightarrow \boldsymbol{y} = A\boldsymbol{x}.$$

(ii) **平移群**(相当于总是取单位矩阵)：

$$\boldsymbol{x} \rightarrow \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{l},$$

与之对应的逆元素是

$$\boldsymbol{y} \rightarrow \boldsymbol{x} = \boldsymbol{y} - \boldsymbol{l}.$$

2.5 正 交 群

2.5.1 正交变换

定义 若一个线性变换

$$\boldsymbol{x} \rightarrow \boldsymbol{y} = A\boldsymbol{x},$$

它在将 $\mathbf{V}_n(\mathbf{R})$中的 $\boldsymbol{x}$ 变换为 $\boldsymbol{y}$ 后有

$$(\boldsymbol{y},\boldsymbol{y}) = y_1^2 + \cdots + y_n^2 = x_1^2 + \cdots + x_n^2 = (\boldsymbol{x},\boldsymbol{x}),$$

则该变换称为 $\mathbf{V}_n(\mathbf{R})$中的**正交变换**.

注意到由$(\boldsymbol{y},\boldsymbol{y})=(\boldsymbol{x},\boldsymbol{x})$的条件可以证明 A 必为非奇异方阵. 为此，我们先引进转置矩阵和正交矩阵的概念.

2.5.2 转置矩阵

定义 对于数域 $\mathbf{F}$ 上的一个 $m\times n$ 矩阵 $A=\{a_{ik}\}$，其行与列的元素互换后所得 $n\times m$ 矩阵 $A^{\mathrm{T}}=\{a_{ki}\}$(即$(A^{\mathrm{T}})_{ik}=a_{ik}^{\mathrm{T}}=a_{ki}$)，称为 A 的**转置矩阵**，有时也记作 $\widetilde{A}$.

设 A 是一个 n 阶方阵，若 $A^{\mathrm{T}}=A$，则 A 称为**对称矩阵**；若 $A^{\mathrm{T}}=-A$，则 A 称为**反称矩阵**.

注意到$(BC)^{\mathrm{T}}=C^{\mathrm{T}}B^{\mathrm{T}}$，可以证明如下：

$$[(BC)^{\mathrm{T}}]_{ij} = (BC)_{ji} = \sum_k b_{jk}c_{ki},$$

$$[C^{\mathrm{T}}B^{\mathrm{T}}]_{ij} = \sum_k (C^{\mathrm{T}})_{ik}(B^{\mathrm{T}})_{kj} = \sum_k c_{ki}b_{jk},$$

左右两边对应矩阵元相等，得证.

2.5.3 标积的定义

根据转置矩阵的定义，对于列矢量 $\boldsymbol{x}$，其转置 $\widetilde{\boldsymbol{x}}$ 为行矢量，即

$$\widetilde{\boldsymbol{x}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}^{\mathrm{T}} = [x_1 \cdots x_n].$$

于是,按矩阵乘法,行矢量 $\tilde{\boldsymbol{x}}$ 与列矢量 $\boldsymbol{y}$ 之积:

$$\tilde{\boldsymbol{x}} \cdot \boldsymbol{y} = [x_1 \cdots x_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + \cdots + x_n y_n = \sum_i x_i y_i$$

是一标量,对于实数域 $\mathbf{R}$ 而言,这种运算称为**标积**. 一般记作$(\boldsymbol{x},\boldsymbol{y})$.

由此可见,

$$(\boldsymbol{x},\boldsymbol{x}) = \tilde{\boldsymbol{x}} \cdot \boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}} \cdot \boldsymbol{x} = \sum_i x_i^2.$$

2.5.4 正交矩阵

定义 若实数域 $\mathbf{R}$ 上的 $n\times n$ 矩阵 R 与其转置矩阵 $\tilde{R}$ 满足下列条件:

$$\tilde{R}R = 1 \quad 即 \quad \tilde{R} = R^{-1},$$

则将 R 称为**正交矩阵**.

于是由正交矩阵所代表的线性变换:

$$\boldsymbol{y} = R\boldsymbol{x}$$

是正交变换,因为

$$(\boldsymbol{y},\boldsymbol{y}) = \tilde{\boldsymbol{y}} \cdot \boldsymbol{y} = \tilde{\boldsymbol{x}}\tilde{R} \cdot R\boldsymbol{x} = \tilde{\boldsymbol{x}} \cdot \boldsymbol{x} = (\boldsymbol{x},\boldsymbol{x}).$$

2.5.5 正交变换保持标积不变

任何两个矢量的标积在正交变换下保持不变.

证 设正交变换 R 使两矢量 $\boldsymbol{x}',\boldsymbol{x}''$变为 $\boldsymbol{y}',\boldsymbol{y}''$,

$$\boldsymbol{x}' \rightarrow \boldsymbol{y}' = R\boldsymbol{x}', \quad \boldsymbol{x}'' \rightarrow \boldsymbol{y}'' = R\boldsymbol{x}'';$$

其标积$(\boldsymbol{x}',\boldsymbol{x}'')$变为$(\boldsymbol{y}',\boldsymbol{y}'')$,由于

$$(\boldsymbol{y}',\boldsymbol{y}'') = \tilde{\boldsymbol{y}}' \cdot \boldsymbol{y}'' = \tilde{\boldsymbol{x}}'\tilde{R} \cdot R\boldsymbol{x}'' = \tilde{\boldsymbol{x}}' \cdot \boldsymbol{x}'' = (\boldsymbol{x}',\boldsymbol{x}'');$$

所以标积保持不变(这里用了$\widetilde{R\boldsymbol{x}'} = \tilde{\boldsymbol{x}}'\tilde{R}$ 和 $\tilde{R}R = 1$ 的性质). 证毕.

2.5.6 等价关系

为了下面的应用,先介绍等价关系的概念.

定义 集合 S 中元素的序对(a,b)有二元关系 $a\sim b$,若满足下列三条件

(i) **自反性** 对任意 a 有 $a\sim a$,

(ii) **对称性** $a\sim b$ 蕴涵 $b\sim a$,

(iii) **传递性** $a\sim b, b\sim c$ 蕴涵 $a\sim c$;

则称之为**等价关系**.

例如,普通用"="相联系的是等价关系,而用">"或"<"相联系的则不是.

2.5.7 正交群

$\mathbf{V}_n(\mathbf{R})$中全体正交变换(或正交矩阵)构成一个群,称为**正交群**,又称**转动反射群**,记作 $\mathbf{O}_n(\mathbf{R})$或 $\mathbf{O}(n)$. $\mathbf{O}_n(\mathbf{R})$中所有行列式为+1 的正交矩阵也构成群,称为**特殊正交群**;又称**转动群**,记作 $\mathbf{SO}_n(\mathbf{R})$或 $\mathbf{SO}(n)$.

现在用$(\boldsymbol{x},\boldsymbol{x})=(\boldsymbol{y},\boldsymbol{y})$是等价关系来证明正交变换构成一个群.

(i) 若 A 和 B 代表正交变换,则

$$\boldsymbol{x}\to\boldsymbol{y}=A\boldsymbol{x}\Longrightarrow(\boldsymbol{y},\boldsymbol{y})=(\boldsymbol{x},\boldsymbol{x}),$$
$$\boldsymbol{y}\to\boldsymbol{z}=B\boldsymbol{y}\Longrightarrow(\boldsymbol{z},\boldsymbol{z})=(\boldsymbol{y},\boldsymbol{y});$$

而先后两次正交变换之积由 BA 代表,结果是

$$\boldsymbol{x}\to\boldsymbol{z}=(BA)\boldsymbol{x}\Longrightarrow(\boldsymbol{z},\boldsymbol{z})=(\boldsymbol{x},\boldsymbol{x});$$

后者是根据等价关系的传递性得到的. 所以,正交变换之积仍是正交变换,满足封闭性条件.

(ii) 满足结合律,因为$(CB)A=C(BA)$,以及

$$(\boldsymbol{t},\boldsymbol{t})=(\boldsymbol{y},\boldsymbol{y}),(\boldsymbol{y},\boldsymbol{y})=(\boldsymbol{x},\boldsymbol{x})$$

和

$$(\boldsymbol{t},\boldsymbol{t})=(\boldsymbol{z},\boldsymbol{z}),(\boldsymbol{z},\boldsymbol{z})=(\boldsymbol{x},\boldsymbol{x})$$

都导致

$$(\boldsymbol{t},\boldsymbol{t})=(\boldsymbol{x},\boldsymbol{x}).$$

(iii) 单位元素是恒同变换 $\boldsymbol{x}\to\boldsymbol{y}=1\boldsymbol{x}=\boldsymbol{x}$,而根据等价关系的自反性有$(\boldsymbol{x},\boldsymbol{x})=(\boldsymbol{x},\boldsymbol{x})$.

(iv) 逆元素就是逆变换 $\boldsymbol{y}\to\boldsymbol{x}=A^{-1}\boldsymbol{y}$,因为 A 是正交矩阵,故 $\widetilde{A}A=1$,即 $A^{-1}=\widetilde{A}$ 存在,同时由于等价关系的对称性,

$$(\boldsymbol{y},\boldsymbol{y})=(\boldsymbol{x},\boldsymbol{x})\Longrightarrow(\boldsymbol{x},\boldsymbol{x})=(\boldsymbol{y},\boldsymbol{y}),$$

所以逆变换存在,它也是正交变换.

这样就证明了,全体正交变换构成正交群. 它是一般线性群中满足正交条件的一个子群.

2.5.8 刚体运动的 Euclid 群

对于仿射群 $\mathbf{Aff}_n(\mathbf{R})$:

$$\boldsymbol{x}\to\boldsymbol{y}=A\boldsymbol{x}+\boldsymbol{l},$$

若将变换矩阵 A 限制为正交矩阵,这样构成的子群称为刚体运动的 **Euclid 群** $\mathbf{E}_n(\mathbf{R})$.

Euclid 群包含正交群作为其子群.

对 $n=3$ 的情况，有 **Euclid 几何学**的基础，后者研究在正交变换 $\mathbf{O}_3(\mathbf{R})$ 下为不变的全部性质.

注 这里所说的**刚体**指的是任意两点之间的距离是不变量，因为对距离的平方有

$$
\begin{aligned}
(\boldsymbol{y}'-\boldsymbol{y}'',\boldsymbol{y}'-\boldsymbol{y}'') &= (A\boldsymbol{x}'-A\boldsymbol{x}'',A\boldsymbol{x}'-A\boldsymbol{x}'') \\
&= (\widetilde{\boldsymbol{x}}'-\widetilde{\boldsymbol{x}}'')\widetilde{A}\cdot A(\boldsymbol{x}'-\boldsymbol{x}'') \\
&= (\widetilde{\boldsymbol{x}}'-\widetilde{\boldsymbol{x}}'')\cdot(\boldsymbol{x}'-\boldsymbol{x}'') \\
&= (\boldsymbol{x}'-\boldsymbol{x}'',\boldsymbol{x}'-\boldsymbol{x}'').
\end{aligned}
$$

另外，Euclid 几何学中定义矢量 $\boldsymbol{x}$ 之**长度**为

$$|\boldsymbol{x}'| = \sqrt{(\boldsymbol{x}',\boldsymbol{x}')},$$

定义两个非零矢量 $\boldsymbol{x}',\boldsymbol{x}''$之间的**夹角**为

$$\theta = \arccos\frac{(\boldsymbol{x}',\boldsymbol{x}'')}{\sqrt{(\boldsymbol{x}',\boldsymbol{x}')(\boldsymbol{x}'',\boldsymbol{x}'')}},$$

而两点 $\boldsymbol{x}',\boldsymbol{x}''$之间的**距离**为

$$d=\sqrt{(\boldsymbol{x}'-\boldsymbol{x}'',\boldsymbol{x}'-\boldsymbol{x}'')};$$

这是普通三维空间的解析几何学中相当概念的自然推广.

2.6 幺 正 群

2.6.1 幺正变换

定义 对于复数域 **C** 上的一个非奇异 n 阶方阵 $A=\{a_{ik}\}$，如果在将矢量空间 $\mathbf{V}_n(\mathbf{C})$ 中的矢量 $\boldsymbol{x}$ 变换为矢量 $\boldsymbol{y}$：

$$\boldsymbol{x}\to\boldsymbol{y}=A\boldsymbol{x}$$

时，能使(用 * 表示复共轭)

$$
\begin{aligned}
(\boldsymbol{y},\boldsymbol{y}) = \widetilde{\boldsymbol{y}}^*\cdot\boldsymbol{y} &= y_1^*y_1+\cdots+y_n^*y_n \\
&= x_1^*x_1+\cdots+x_n^*x_n = \widetilde{\boldsymbol{x}}^*\cdot\boldsymbol{x}=(\boldsymbol{x},\boldsymbol{x})
\end{aligned}
$$

保持不变，则该变换称为**幺正变换**.

2.6.2 Hermite 矩阵

定义 对于复数域 **C** 上的一个 $m\times n$ 矩阵 $A=\{a_{jk}\}$，其行与列的元素互换并取复共轭(将 i 变为 −i)后所得($n\times m$)矩阵 $A^\dagger=\{a_{kj}^*\}$(即 $(A^\dagger)_{jk}=\widetilde{a}_{jk}^*=a_{kj}^*$，这里 $a^\dagger=\widetilde{A}^*$，* 表示复共轭，～表示转置，上标 † 表示 Hermite 共轭(复共轭转置))称为 A 的 **Hermite 共轭矩阵**，又称**复共轭转置矩阵**.

对于 n 阶方阵 A，若满足条件 $A^{\dagger}=A$，则称为 **Hermite 矩阵.**

对于 Hermite 共轭矩阵，同样可以证明

$$(BC)^{\dagger} = C^{\dagger}B^{\dagger}.$$

2.6.3 幺正矩阵

定义 若复数域 **C** 上的 $n\times n$ 矩阵 U 与其 Hermite 共轭矩阵 $U^{\dagger}$ 满足下列条件

$$U^{\dagger}U = 1 \quad 即 \quad U^{\dagger} = U^{-1},$$

则将 U 称为**幺正矩阵.**

于是，由幺正矩阵 U 所代表的线性变换

$$\boldsymbol{y} = U\boldsymbol{x}$$

是幺正变换，因为

$$(\boldsymbol{y},\boldsymbol{y}) = \boldsymbol{y}^{\dagger}\cdot\boldsymbol{y} = \boldsymbol{x}^{\dagger}U^{\dagger}\cdot U\boldsymbol{x} = \boldsymbol{x}^{\dagger}\cdot\boldsymbol{x} = (\boldsymbol{x},\boldsymbol{x}).$$

2.6.4 幺正变换保持标积不变

对于复数域 **C** 上的列矢量 $\boldsymbol{x}$，其 Hermite 共轭矢量 $\boldsymbol{x}^{\dagger}=\boldsymbol{x}^{*\mathrm{T}}=\widetilde{\boldsymbol{x}}^{*}$ 为行矢量

$$\boldsymbol{x}^{\dagger} = \widetilde{\boldsymbol{x}}^{*} = [x_1^{*},\cdots,x_n^{*}].$$

相应地，标积的定义修正为

$$(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x}^{\dagger}\cdot\boldsymbol{y} = x_1^{*}y_1 + \cdots + x_n^{*}y_n = \sum_i x_i^{*}y_i.$$

由此可见，幺正变换保持标积不变，因为

$$(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x}^{\dagger}\cdot\boldsymbol{y} = \boldsymbol{x}^{\dagger}U^{\dagger}\cdot U\boldsymbol{y} = (U\boldsymbol{x})^{\dagger}\cdot(U\boldsymbol{y}) = (U\boldsymbol{x},U\boldsymbol{y}).$$

另一方面，

$$(\boldsymbol{x},\boldsymbol{x}) = \boldsymbol{x}^{\dagger}\cdot\boldsymbol{x} = x_1^{*}x_1 + \cdots + x_n^{*}x_n = |x_1|^2 + \cdots + |x_n|^2 = \sum_i |x_i|^2$$

为一实数.

2.6.5 幺正群

$\mathbf{V}_n(\mathbf{C})$ 中全体幺正变换（或幺正矩阵）构成一个群，称为**幺正群**，记作 $\mathbf{U}_n(\mathbf{C})$ 或 $\mathbf{U}(n)$. $\mathbf{U}_n(\mathbf{C})$ 中所有行列式为 $+1$ 的幺正矩阵也构成群，称为**特殊幺正群**，记作 $\mathbf{SU}_n(\mathbf{C})$ 或 $\mathbf{SU}(n)$.

幺正群 $\mathbf{U}_n(\mathbf{C})$ 是一般线性群 $\mathbf{GL}_n(\mathbf{C})$ 的子群.

以上结果可利用 $(\boldsymbol{x},\boldsymbol{x})=(\boldsymbol{y},\boldsymbol{y})$ 是一个等价关系，与 2.5.7 节完全类似地予以证明，这里从略.

2.7 置 换 群

2.7.1 置换的定义

定义 对于变换

$$x \to y = Px,$$

若其结果 $y_1, y_2, \cdots, y_n$ 只不过是 $x_1, x_2, \cdots, x_n$ 的一种新排列,而其次序可能与原次序不同,这种变换称为 n 元**置换**.

或者说,n 元置换是 n 元集合到其自身的一个一一映射.

2.7.2 置换矩阵

定义 若是一个 $n \times n$ 矩阵的每行中和每列中只有一个矩阵元为 1(其余均为 0),总共出现 n 个 1;这样的矩阵称为**置换矩阵**.

例如:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_1 \\ x_2 \end{bmatrix}.$$

显然,置换矩阵是正交矩阵,证明如下:

若置换矩阵为 $P = \{p_{ij}\}$,则有

$$(\tilde{P}P)_{ij} = \sum_k p_{ki} p_{kj} = \delta_{ij} \sum_k (p_{ki})^2 = \delta_{ij};$$

因为,k 行中只可能在一列处不为零,故而有 $p_{ki} p_{kj} = \delta_{ij} (p_{ki})^2$,$\delta_{ij} = 1 (i = j)$ 或 $0 (i \neq j)$ 是 Kronecker 符号. 又因为 i 列中必然会在某一行处为 1 而其余处为零,故有 $\sum_k (p_{ki})^2 = 1$;所以

$$\tilde{P}P = 1,$$

即 P 是正交矩阵.

2.7.3 对称群的定义

(1) $\mathbf{V}_n(\mathbf{R})$中的全体置换构成正交群 $\mathbf{O}_n(\mathbf{R})$的一个子群,称为**对称变换群**,记作 $\mathbf{S}_n(\mathbf{R})$.

证 因为相继两次置换(乘法)的结果仍是 $x_1, \cdots, x_n$ 的一种排列,显然代表一种置换的结果(群中元素),这是封闭性. 由于置换矩阵的运算满足结合律,有单位矩阵 I 和逆矩阵 $P^{-1} = \tilde{P}$,满足群定义的后三个条件. 所以,全体置换构成群.

至于 $\mathbf{S}_n(\mathbf{R})$ 是 $\mathbf{O}_n(\mathbf{R})$ 的子群则由置换矩阵属于正交矩阵而得证.

(2) 特别是,对于任意 n 个文字 $1,2,\cdots,n$ 的全体置换构成一个群,称为 n 个文字的**对称群** $\mathbf{S}_n$. $\mathbf{S}_n$ 与 $\mathbf{S}_n(\mathbf{R})$ 同构.

证 n 元集合 $x_1,x_2,\cdots,x_n$ 与 n 元集合 $1,2,\cdots,n$ 之间可建立一一对应关系:

$$1\leftrightarrow x_1,\quad 2\leftrightarrow x_2,\quad \cdots,\quad n\leftrightarrow x_n.$$

于是,显然 n 个文字 $1,2,\cdots,n$ 的全体置换构成群;并且有

$$P_a\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\leftrightarrow P_a\begin{bmatrix}1\\ \vdots\\ n\end{bmatrix},\quad P_b\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\leftrightarrow P_b\begin{bmatrix}1\\ \vdots\\ n\end{bmatrix},$$

$$P_aP_b\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\leftrightarrow P_aP_b\begin{bmatrix}1\\ \vdots\\ n\end{bmatrix},$$

即两群元素之间有一一对应关系,所以它们是同构的.

(3) 对称群是**有限群**,其阶为 $n!$

证 n 元置换总共有 $n!$ 个,这可由置换矩阵的构成予以证明,因为

第 1 行的 1 可于 n 列中选定任一处,

第 2 行的 1 可于其余 $n-1$ 列中选定任一处,

$\vdots$

第 i 行的 1 可于其余 $n-i+1$ 列中选定任一处,

$\vdots$

第 $n-1$ 行的 1 可于其余 2 列中选定一处,

第 n 行的 1 只剩下 1 种可能;

即共有 $n(n-1)\cdots(n-i+1)\cdots2\cdot1=n!$ 个不同的置换矩阵,也就是说,总共有 $n!$ 个不同的置换结果. 所以,n 个文字的对称群的元素总共只有 $n!$ 个;故对称群为**有限群**,而群元素总数 $n!$ 为其阶.

2.7.4 置换、轮换与对换

(1) 置换的记号

首先引进**置换的记号**. n 个文字的集合

$$\mathrm{I}_n=\{1,2,\cdots,n\},$$

经置换后,若 $1\to p_1,2\to p_2,\cdots,n\to p_n$,则记为

$$1\ \ 2\ \ \cdots\ \ n\to p_1\ \ p_2\ \ \cdots\ \ p_n\quad 或\quad \begin{pmatrix}1 & 2 & \cdots & n\\ p_1 & p_2 & \cdots & p_n\end{pmatrix};$$

这里 $1\ \ 2\ \ \cdots\ \ n$ 为按自然序或标准序的排列,而置换后的结果 $p_1\ \ p_2\ \ \cdots\ \ p_n$ 为不

同的有序排列；若只强调结果，还可简单记为

$$p_1 \quad p_2 \quad \cdots \quad p_n.$$

(2) 轮换①

集合 $\mathrm{I}_n=\{1,2,\cdots,n\}$ 的一个置换 γ，它将集合中的元素序列 $a_1,a_2,\cdots,a_r(1<r\leqslant n)$ 循环地予以置换，使

$$\gamma(a_1)=a_2,\quad \gamma(a_2)=a_3,\quad \cdots,\quad \gamma(a_{r-1})=a_r,\quad \gamma(a_r)=a_1,$$

而 I_n 中的其余数保持不变，这样的置换称为 r 元**轮换**，记作：

$$\gamma=(a_1a_2\cdots a_r).$$

显然，同样可以记作：

$$\gamma=(a_2a_3\cdots a_ra_1)=(a_3a_4\cdots a_ra_1a_2)=\cdots=(a_ra_1\cdots a_{r-1}).$$

而且，置换 γ^2 会使

$$a_1\to a_3,\quad a_2\to a_4,\quad \cdots,\quad a_r\to a_2;$$

以及，一般地，对 $1\leqslant k\leqslant r$，

$$\gamma^k(a_i)=a_{i+k},\quad 若\ i+k\leqslant r,$$
$$\gamma^k(a_i)=a_{i+k-r},\quad 若\ i+k>r.$$

这显然表明 $\gamma^r=1$，而若 $1\leqslant k<r$，则 $\gamma^k\neq 1$. 因此，γ 是 r 阶.

两个轮换 γ 和 γ'，若它们所施作用序列无公共元素，则称为**不相交的**. 可以看出，两个不相交轮换可交换，即 $\gamma\gamma'=\gamma'\gamma$.

设置换 σ 是若干不相交轮换之积：

$$\sigma=(a_1a_2\cdots a_r)(b_1b_2\cdots b_s)\cdots(c_1c_2\cdots c_t),$$

若 m 是 $r,s,\cdots,t$ 的最小公倍数，则容易证明 σ 的阶是 m.

若令记号 (j) 表示 $j\to j$ 并称为“1 元轮换”. 则有置换的下述轮换表示法.

(3) 置换的轮换表示法②

定理 任一置换可表示为不相交轮换之积.

例如：

$$\sigma=\begin{pmatrix}1&2&3&4&5&6&7&8\\3&6&5&4&8&2&7&1\end{pmatrix}$$
$$=(1\ 3\ 5\ 8)(2\ 6)(4)(7).$$

一般地，对于任一 σ，可从 $\mathrm{I}_n=\{1,2,\cdots,n\}$ 中任一数字（比如 a_1）开始，并构成 $\sigma(a_1)=a_2,\sigma(a_2)=a_3,\cdots$，直到 $\sigma(a_r)=a_1$；这样得到一个 r 元轮换 $(a_1,a_2,\cdots,a_r)$. 若 $r<n$，在其余数字中选一 b_1，类似地得到一个 s 元轮换 $(b_1,\cdots,b_s)$. 再在剩余数字中选择并继续下去，直到穷举完集合 I_n 中的数字为止. 这样就将置换 σ 分解为不相

---

① 参考：N. Jacobson, *Basic Algebra* I, Freeman and Company, San Francisco, 1974, pp. 48—50.

② 参考上注 N. Jacobson 书.

交轮换之积：

$$\sigma = (a_1, \cdots, a_r)(b_1, \cdots, b_s)\cdots(c_1, \cdots, c_t);$$

并且容易证明，这种分解是唯一的. 当然，应注意到在此分解中不同轮换之间次序可交换，而且对于平凡的“1 元轮换”可以将之包含在内，也可予以省略. 这种分解称为**置换的轮换表示法**.

(4) 对换

对于 $\mathbf{S}_n$ 中的一个置换 τ，即 $\tau\in\mathbf{S}_n$，存在一个整数 i 使得 $1\leqslant i\leqslant n-1$ 和

$$\tau(i) = i+1, \tau(i+1) = i, \text{以及 } \tau(k) = k(\text{若 } k \neq i, i+1),$$

则此置换称为对换，并记作 τ_i 或 $(i, i+1)$.

对于更一般的对换 $\tau(\tau\in\mathbf{S}_n)$ 可定义为：若存在一对整数 k, l 使得 $1\leqslant k<l\leqslant n$ 和

$$\tau(i) = i(\text{若 } i \neq k, l), \quad \tau(k) = l, \quad \tau(l) = k;$$

则称为 k, l 之间的**对换**，记作 $(k\ l)$.

2.7.5　对称群有关定理

下面介绍几个与对称群有关的定理.

(1) 重排定理

定理　对称群 $\mathbf{S}_n$ 中所有元素(总共 $n!$ 个元素)乘以群中任一特定元素所得之积重现原对称群中所有元素.

证　设 $\sigma_p, \sigma_i, \sigma_j\in\mathbf{S}_n$，则用 σ_p 乘 σ_i, σ_j 后，必有

(i) $\sigma_p\sigma_i$ 和 $\sigma_p\sigma_j$ 仍为 $\mathbf{S}_n$ 中元素，

(ii) $\sigma_p\sigma_i\neq\sigma_p\sigma_j$，若 $\sigma_i\neq\sigma_j$.

(i) 是由于群元素的封闭性. 至于(ii)，如果不然，假定它们变为 $\mathbf{S}_n$ 中的同一元素：

$$\sigma_p\sigma_i = \sigma_p\sigma_j,$$

则再用 σ_p^{-1} 对同一元素作用的结果(由结合律)会得出

$$\sigma_i = \sigma_j$$

与原条件 $\sigma_i\neq\sigma_j$ 矛盾，因而 $\sigma_p\sigma_i=\sigma_p\sigma_j$ 这一假定是错误的，所以(ii)的结论是正确的. 由此可见，用任一 σ_p 乘以 $\mathbf{S}_n$ 中总共 $n!$ 个不同元素(包括 σ_p)会得 $\mathbf{S}_n$ 中的全部 $n!$ 个不同元素，即重现原对称群. 证毕.

应该注意到，以上结果对一切有限群均适用，称为**重排定理**.

(2) 对称群中元素均可由对换生成[①]

定理　对称群 $\mathbf{S}_n$ 中全部群元素均可由其中所含对换予以生成.

可采用归纳法证明任一置换 $\sigma\in\mathbf{S}_n$ 是若干对换之积.

$n=1$ 时只有 1 个元素 1 和 $n=2$ 时有 2 个元素 12 和 21，定理显然成立. 若定

---

① 参考：R. Godement，*Cours d'Algèbre*，Hermann，Paris，1963.(英译本，1969，p. 130.)

理对 S_{n-1} 成立，要证明对 S_n 也成立.

考虑任一 $\sigma\in S_n$，并令 $\sigma(n)=i$，即元素 σ 中第 n 个文字为 i. 若 τ_j 表示 j 和 $j+1$ 对换，则显然下列置换(共含 $n-i$ 个 τ 因子)：

$$\mu=\tau_{n-1}\circ\cdots\circ\tau_i\circ\sigma$$

满足 $\mu(n)=n$，由于 $\tau^{-1}=\tau$，于是有

$$\sigma=\tau_i^{-1}\circ\cdots\circ\tau_{n-1}^{-1}\circ\mu=\tau_i\circ\cdots\circ\tau_{n-1}\circ\mu.$$

因此，为了证明 σ 是对换之积，只要证明 μ 是对换之积就可以了. 但因 $\mu(n)=n$，μ 对 I_n 的 $1,2,\cdots,n-1$ 进行置换；换句话说，μ“归结为”作用于 I_{n-1} 上的一个置换 $\mu'\in S_{n-1}$. 根据归纳假设，μ' 可写成

$$\mu'=\nu_1\circ\cdots\circ\nu_q,$$

其中 $\nu_1,\cdots,\nu_q$ 是群 S_{n-1} 中的对换. 若定义

$$\omega_j(x)=\begin{cases}\nu_j(x), & 对\ x\in\{1,2,\cdots,n-1\},\\ n, & 对\ x=n;\end{cases}$$

由于 μ 和 μ' 在集合 I_{n-1} 上一致且 $\mu(n)=n$，显然有

$$\mu=\omega_1\circ\cdots\circ\omega_q.$$

但因 ν_j 是 I_{n-1} 上的对换，ω_j 显然是 I_n 上的对换. 因此，μ 是对称群 S_n 中对换之积，所以 σ 也是如此. 证毕.

(3) 对称群中奇偶置换各半

首先定义**置换的奇偶性**.

定义 一个置换表示成对换之积时，若对换个数为奇数，称为**奇置换**；若对换个数为偶数，称为**偶置换**.

可以证明，对称群中奇偶置换各占一半，即有下述定理.

定理 对称群 S_n 的全部 $n!$ 个置换中，有 $\frac{1}{2}n!$ 个是偶置换和 $\frac{1}{2}n!$ 个是奇置换.

证 用归谬法证明. 如果定理不真，假定偶置换数 N_2 与奇置换数 N_1 不等：$N_2\neq N_1$. 用一个对换 $\tau_i=(i,i+1)$ 乘 S_n 中全部元素后，由重排定理知仍得原对称群，但由定义知现在偶置换数变为 N_1 而奇置换数变为 N_2(因为对换个数增加 1 使偶变奇和奇变偶)，产生矛盾. 假定 $N_2\neq N_1$ 为谬，即应有 $N_2=N_1=\frac{1}{2}n!$，定理得证.

2.7.6 置换群

对称群 S_n 的子群通称**置换群**.

(1) 交错群

全体偶置换构成 S_n 的一个子群，称为**交错群**，记作 A_n；其阶为 $\frac{1}{2}n!$.

这是因为两个偶置换之积为偶置换，恒同置换为偶置换，偶置换之逆为偶置

换，故均为群中元素；而偶置换总数为$\frac{1}{2}n!$.

注 全体奇置换不能构成群，因为两个奇置换之积为偶置换，恒同置换为偶置换，均非集合中元素.

(2) 循环群

若G中有一元素α使$\langle\alpha\rangle=G$，这里用$\langle\alpha\rangle$表示所有α^m的集合(m为整数)，则G称为**循环群**；若有最小正整数r使$\alpha^r=1$(1为单位元素)，则称为r阶循环群. α称为G的**生成元**. 因此，循环群是具有单个生成元的群，所有循环群都是**Abel群**.

以上是其抽象定义. 因为$\alpha^m\cdot\alpha^l=\alpha^{m+l}$也是$\langle\alpha\rangle$中元素；而$\alpha^0=1$；$(\alpha^m)^{-1}=(\alpha^{-1})^m=\alpha^{-m}$亦均为$\langle\alpha\rangle$中元素；另外，结合律成立也是很明显的；所以$\langle\alpha\rangle$的确构成群；而且由于$\alpha^m\cdot\alpha^l=\alpha^l\cdot\alpha^m$，故为Abel群.

具体地说，对称群$\mathbf{S}_n$中有许多这样的循环群，它们都是有限群. 例如：

由r元轮换

$$\gamma=(1\ 2\ \cdots\ r)$$

作为生成元构成的集合

$$\langle\gamma\rangle=\{1,\gamma,\gamma^2,\cdots,\gamma^{r-1}\}$$

构成一循环群，其阶为r.

由2.7.4节(3)中给出的置换

$$\sigma=(a_1\cdots a_r)(b_1\cdots b_s)\cdots(c_1\cdots c_t)$$

作为生成元构成的集合

$$\langle\sigma\rangle=\{1,\sigma,\sigma^2,\cdots,\sigma^{m-1}\}$$

构成一循环群，其阶为m，这里m是$r,s,\cdots,t$的最小公倍数.

以及诸如此类等等.

2.8 群同构的具体例子

(1) 3次对称群 $\mathbf{S}_3$

3个文字1,2,3的对称群$\mathbf{S}_3$，其阶为3!，总共包括6个元素，它们是

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},$$
$$\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix},\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}.$$

与$\mathbf{S}_3$同构的有：

3次对称变换群$\mathbf{S}_3(\mathbf{R})$，见表2.2中前两行；

3边的二面体群$\mathbf{D}_3$，见表2.2中后两行.

表 2.2 S_3 群元素记号及与同构群的对比

$S_3(R)$	结果	$x_1x_2x_3$	$x_3x_1x_2$	$x_2x_3x_1$	$x_1x_3x_2$	$x_3x_2x_1$	$x_2x_1x_3$
	记号	$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$	$\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}$	$\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$	$\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$	$\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}$	$\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}$
S_3	结果	123	312	231	132	321	213
	记号	$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}$	$\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$	$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$	$\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$	$\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$	$\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$
		$\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}2\\2\end{pmatrix}\begin{pmatrix}3\\3\end{pmatrix}$	$\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$	$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$	$\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}2&3\\3&2\end{pmatrix}$	$\begin{pmatrix}2\\2\end{pmatrix}\begin{pmatrix}1&3\\3&1\end{pmatrix}$	$\begin{pmatrix}3\\3\end{pmatrix}\begin{pmatrix}1&2\\2&1\end{pmatrix}$
		1	(321)	(123)	(23)	(13)	(12)
		(1)(2)(3)	(321)	(123)	(1)(23)	(2)(13)	(3)(12)
	分解	(321)(123)		$(321)^2$	(321)(12)	(321)(23)	(321)(13)
		(123)(321)	$(123)^2$		(123)(13)	(123)(12)	(123)(23)
		$(23)^2$	(23)(12)	(23)(13)		(23)(123)	(23)(321)
		$(13)^2$	(13)(23)	(13)(12)	(13)(321)		(13)(123)
		$(12)^2$	(12)(13)	(12)(23)	(12)(123)	(12)(321)	
D_3	结果	A; B C	C; A B	B; C A	A; C B	C; B A	B; A C
	记号	e	$R_1(R)$	$R_2(R^2)$	$R_A(S)$	$R_B(SR^2)$	$R_C(SR)$

1. 注意到对于轮换有$(ij)=(ji)$,$(ijk)=(jki)=(kij)$.
2. 框中记号补足了上行记号中省略掉的“1 元轮换”,以示完全.

关于对称群 $\mathbf{S}_3$ 中元素的各种记号和分解情况，以及与同构群的对比，均见表 2.2.

关于二面体群 $\mathbf{D}_3$ 与对称群 $\mathbf{S}_3$ 的乘法表，分别见表 2.3 与表 2.4.

表 2.3 $\mathbf{D}_3$ 群乘法表

	e	R_1	R_2	R_A	R_B	R_C
e	e	R_1	R_2	R_A	R_B	R_C
R_2	R_2	e	R_1	R_C	R_A	R_B
R_1	R_1	R_2	e	R_B	R_C	R_A
R_A	R_A	R_C	R_B	e	R_2	R_1
R_B	R_B	R_A	R_C	R_1	e	R_2
R_C	R_C	R_B	R_A	R_2	R_1	e

表 2.4 $\mathbf{S}_3$ 群乘法表

	1	(321)	(123)	(23)	(13)	(12)
1	1	(321)	(123)	(23)	(13)	(12)
(123)	(123)	1	(321)	(12)	(23)	(13)
(321)	(321)	(123)	1	(13)	(12)	(23)
(23)	(23)	(12)	(13)	1	(123)	(321)
(13)	(13)	(23)	(12)	(321)	1	(123)
(12)	(12)	(13)	(23)	(123)	(321)	1

顺便提一下，$\mathbf{S}_3$（和 $\mathbf{D}_3$）是具有两个**生成元** a_1,a_2 和**定义关系** $R(a_1,a_2)=e$ 的**有限生成群**：

$$G=\{a_1,a_2\mid R(a_1,a_2)=e\};$$

例如，对 $\mathbf{S}_3$ 可取 $a_1=(321)$，$a_2=(23)$；对 $\mathbf{D}_3$ 可取 $a_1=R$，$a_2=S$；而定义关系则可写成

$$a_1^3=a_2^2=(a_1a_2)^2=e.$$

(2) 3 阶循环群

3 阶循环群的生成元可取 $\alpha=(321)$，则

$$\langle\alpha\rangle=\{1,(321),(321)^2\}=\{1,(321),(123)\},$$

3 个群元素 1，(321)，(123) 正好是 $\mathbf{S}_3$ 中的 3 个偶置换，所以在 ($n=3$) 这个特例下它同时是交错群 $\mathbf{A}_3$，是 $\mathbf{S}_3$ 的子群.

与 3 阶循环群同构的有：

3 边二面体群 $\mathbf{D}_3$ 的子群，绕对称心的转动群，包括 3 个元素 e,R_1,R_2，生成元可取 R_1.

（复）3 次单位根的群 $\mathbf{C}(3)$，它由 1 的 3 个立方根

$$1,\quad \omega,\quad \omega^2\quad \left[\omega = e^{i\frac{2\pi}{3}} = \frac{1}{2}(-1 + i\sqrt{3})\right]$$

构成，生成元可取 ω.

它们的乘法表见表 2.5. 还可注意到，前两表与(1)小节中相应表完全一致.

表 2.5　群乘法表

i. 3 阶循环群

	1	(321)	(123)
1	1	(321)	(123)
(123)	(123)	1	(321)
(321)	(321)	(123)	1

ii. D_3 的子群

	e	R_1	R_2
e	e	R_1	R_2
R_2	R_2	e	R_1
R_1	R_1	R_2	e

iii. C(3)群

	1	ω	ω^2
1	1	ω	ω^2
ω^2	ω^2	1	ω
ω	ω	ω^2	1

只要注意到它们的元素之间的下列一一对应关系：

	循环群		D_3 的子群		C(3)群
α^3：	1	$\longleftrightarrow$	e	$\longleftrightarrow$	1
α：	(321)	$\longleftrightarrow$	R_1	$\longleftrightarrow$	ω
α^2：	(123)	$\longleftrightarrow$	R_2	$\longleftrightarrow$	ω^2

以及相应的乘法表，立即得出它们是同构的. 它们都可记作

$$\mathbf{C}_3 = \{\alpha \mid \alpha^3 = e\}.$$

第 3 章　行　列　式

3.1　行列式的定义

首先给出熟知的 2 阶和 3 阶行列式的例子：

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$

$$- a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}.$$

这里将进一步介绍行列式的定义、性质、运算及应用方面的有关内容. 首先给出行列式的定义.

定义　n 阶矩阵

$$A = \{a_{ij}\} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

的**行列式**记作下列形式之一：

$$\det A = |A| = |a_{ij}| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix},$$

它由下面的展开式

$$|A| = \sum_{n!\text{个置换}} \pm \begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} a_{11}a_{22}\cdots a_{nn}$$

予以定义，这里 $\begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}$ 是作用于后面列标的置换，其前的 $\pm$ 号代表该置换的奇偶性：

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} = \begin{cases} + & \text{对}\genfrac{}{}{0pt}{}{\text{偶}}{\text{奇}}\text{置换}, \\ - & \end{cases}$$

而求和号 $\sum\limits_{n!\text{个置换}}$ 是对全体 $n!$ 个可能的置换求和，或者简单直接地写成下列定义式$\left[\text{这里}\pm=\operatorname{sgn}\begin{pmatrix}1 & 2 & \cdots & n\\ p_1 & p_2 & \cdots & p_n\end{pmatrix}\right]$：

$$|A|=\sum\pm a_{1p_1}a_{2p_2}\cdots a_{np_n}.\qquad(\text{I})$$

现在将行列式中每一项 $a_{1p_1}a_{2p_2}\cdots a_{np_n}$ 的因子按 $p_1p_2\cdots p_n$ 的自然序重新排列，其结果比如说是 $a_{q_1 1}a_{q_2 2}\cdots a_{q_n n}$. 这就是说，

$$\begin{pmatrix}1 & 2 & \cdots & n\\ p_1 & p_2 & \cdots & p_n\end{pmatrix}=\begin{pmatrix}q_1 & q_2 & \cdots & q_n\\ 1 & 2 & \cdots & n\end{pmatrix}$$

为同一置换. 于是，行列式同样可定义为

$$|A|=\sum_{n!\text{个置换}}\pm\begin{pmatrix}q_1 & q_2 & \cdots & q_n\\ 1 & 2 & \cdots & n\end{pmatrix}a_{11}a_{22}\cdots a_{nn},$$

这里$\begin{pmatrix}q_1 & q_2 & \cdots & q_n\\ 1 & 2 & \cdots & n\end{pmatrix}$现在是作用于后面行标上的置换，也可简单直接地写成下列等价定义式$\left[\text{这里}\pm=\operatorname{sgn}\begin{pmatrix}q_1 & q_2 & \cdots & q_n\\ 1 & 2 & \cdots & n\end{pmatrix}\right]$：

$$|A|=\sum\pm a_{q_1 1}a_{q_2 2}\cdots a_{q_n n}.\qquad(\text{II})$$

由于以上情况，(Ⅰ)和(Ⅱ)是行列式的等价定义，可用更简洁记号来表述：

$$|A|=\sum_P(-)^P P[a_{11}a_{22}\cdots a_{nn}],$$

这里 P(最后一个)表示作用于后面列标(或行标)的一个置换，$(-)^P=\operatorname{sgn}P$(对偶置换为"+"而对奇置换为"−")，$\sum\limits_P$ 则是对全体 $n!$ 个可能的置换求和.

3.2 行列式的主要性质

行列式具有以下几个**主要性质**：

(1) 矩阵 A 与转置矩阵 $\widetilde{A}$，其行列式相等.

证 令 $A=\{a_{ij}\}$，$\widetilde{A}=B=\{b_{ij}\}$，有 $b_{ij}=a_{ji}$；根据行列式的等效定义(Ⅰ)和(Ⅱ)，

$$\begin{aligned}|\widetilde{A}|&\xlongequal{(\text{II})}\sum\pm b_{p_1 1}b_{p_2 2}\cdots b_{p_n n}\\ &=\sum\pm a_{1p_1}a_{2p_2}\cdots a_{np_n}\xlongequal{(\text{I})}|A|.\end{aligned}$$

这里还用了

$$\operatorname{sgn}\begin{pmatrix}1 & 2 & \cdots & n\\ p_1 & p_2 & \cdots & p_n\end{pmatrix}=\operatorname{sgn}\begin{pmatrix}p_1 & p_2 & \cdots & p_n\\ 1 & 2 & \cdots & n\end{pmatrix},$$

因为置换与逆置换具有相同的奇偶性.

(2) 若 $A=\{a_{ij}\}$ 和 $B=\{b_{ij}\}=A(r\text{ 列}\leftrightarrow s\text{ 列})$，则其行列式仅相差符号，即 $|B|=-|A|$；对于 $B=A(r\text{ 行}\leftrightarrow s\text{ 行})$ 的情况亦有相同结果.

证　令 $A=(\boldsymbol{a}^{(1)},\cdots,\boldsymbol{a}^{(r)},\cdots,\boldsymbol{a}^{(s)},\cdots,\boldsymbol{a}^{(n)})$,

$B=(\boldsymbol{b}^{(1)},\cdots,\boldsymbol{b}^{(r)},\cdots,\boldsymbol{b}^{(s)},\cdots,\boldsymbol{b}^{(n)})$,

其中 $\boldsymbol{b}^{(r)}=\boldsymbol{a}^{(s)},\boldsymbol{b}^{(s)}=\boldsymbol{a}^{(r)},\boldsymbol{b}^{(t)}=\boldsymbol{a}^{(t)}(t\neq r,t\neq s)$，于是

$$|B|=\sum\operatorname{sgn}\begin{pmatrix}q_1\cdots q_r\cdots q_s\cdots q_n\\1\cdots r\cdots s\cdots n\end{pmatrix}b_{q_11}\cdots b_{q_rr}\cdots b_{q_ss}\cdots b_{q_nn}$$

$$=\sum\operatorname{sgn}\begin{pmatrix}q_1\cdots q_r\cdots q_s\cdots q_n\\1\cdots r\cdots s\cdots n\end{pmatrix}a_{q_11}\cdots a_{q_rs}\cdots a_{q_sr}\cdots a_{q_nn}.$$

由于

$$\begin{pmatrix}q_1\cdots q_r\cdots q_s\cdots q_n\\1\cdots r\cdots s\cdots n\end{pmatrix}=\begin{bmatrix}q_1\cdots q_r\cdots q_s\cdots q_n\\q_1\cdots q_s\cdots q_r\cdots q_n\end{bmatrix}\begin{pmatrix}q_1\cdots q_s\cdots q_r\cdots q_n\\1\cdots r\cdots s\cdots n\end{pmatrix}$$

$$=(q_sq_r)\begin{pmatrix}q_1\cdots q_s\cdots q_r\cdots q_n\\1\cdots r\cdots s\cdots n\end{pmatrix},$$

而对于对换有 $\operatorname{sgn}(q_sq_r)=\operatorname{sgn}(sr)=-$，即

$$\operatorname{sgn}\begin{pmatrix}q_1\cdots q_r\cdots q_s\cdots q_n\\1\cdots r\cdots s\cdots n\end{pmatrix}=-\operatorname{sgn}\begin{pmatrix}q_1\cdots q_s\cdots q_r\cdots q_n\\1\cdots r\cdots s\cdots n\end{pmatrix},$$

以及

$$a_{q_11}\cdots a_{q_rs}\cdots a_{q_sr}\cdots a_{q_nn}=a_{q_11}\cdots a_{q_sr}\cdots a_{q_rs}\cdots a_{q_nn};$$

所以,

$$|B|=-\sum\operatorname{sgn}\begin{pmatrix}q_1\cdots q_s\cdots q_r\cdots q_n\\1\cdots r\cdots s\cdots n\end{pmatrix}a_{q_11}\cdots a_{q_sr}\cdots a_{q_rs}\cdots a_{q_nn}$$

$$=-|A|.$$

证毕.

对于 $B=A(r\text{ 行}\leftrightarrow s\text{ 行})$ 的情况，可以类似地加以证明.

(3) 若矩阵 A 有 (r,s) 二列（或二行）相同，则其行列式为零，即 $|A|=0$.

证　令 $B=A(r\leftrightarrow s)$，则由性质(2)有

$$|B|=-|A|.$$

又由前提条件 $A(r\leftrightarrow s)=A$ 有

$$|B|=|A|,$$

所以结果是 $|A|=0$.

(4) 若 $A=[\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)}]$ 和 $B=[\boldsymbol{b}',\boldsymbol{b}'',\cdots,\boldsymbol{b}^{(n)}]$，且有

$$\boldsymbol{b}^{(r)}=\lambda\boldsymbol{a}^{(r)},\text{而 }\boldsymbol{b}^{(s)}=\boldsymbol{a}^{(s)}(\text{所有 }s\neq r),$$

则

$$|B| = \lambda |A|.$$

证 显然有

$$|B| = \sum \pm b_{q_1 1} \cdots b_{q_r r} \cdots b_{q_n n}$$
$$= \sum \pm a_{q_1 1} \cdots (\lambda a_{q_r r}) \cdots a_{q_n n}$$
$$= \lambda \sum \pm a_{q_1 1} \cdots a_{q_r r} \cdots a_{q_n n} = \lambda |A|.$$

注 若 $B=\lambda A$,则有 $|B|=\lambda^n |A|$.

(5) 若矩阵 A 中二列(或二行)元素成比例,则其行列式之值为零.

证 提出比例常数[性质(4)]后二列(或二行)相同,故行列式为零[性质(3)],得证.

(6) 若 $A=[\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)}]$,$B=[\boldsymbol{b}',\boldsymbol{b}'',\cdots,\boldsymbol{b}^{(n)}]$和 $C=[\boldsymbol{c}',\boldsymbol{c}'',\cdots,\boldsymbol{c}^{(n)}]$具有

$$\boldsymbol{c}^{(i)} = \boldsymbol{b}^{(i)} = \boldsymbol{a}^{(i)}\text{(所有 } i \neq r\text{),而 } \boldsymbol{c}^{(r)} = \boldsymbol{a}^{(r)} + \boldsymbol{b}^{(r)},$$

则

$$|C| = |A| + |B|.$$

证

$$|C| = \sum \pm a_{q_1 1} \cdots (a_{q_r r} + b_{q_r r}) \cdots a_{q_n n}$$
$$= \sum \pm a_{q_1 1} \cdots a_{q_r r} \cdots a_{q_n n} + \sum \pm b_{q_1 1} \cdots b_{q_r r} \cdots b_{q_n n}$$
$$= |A| + |B|.$$

注 对于行矢量的情况有类似结果.

(7) 若 $A=[\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)}]$和 $B=[\boldsymbol{b}',\boldsymbol{b}'',\cdots,\boldsymbol{b}^{(n)}]$具有

$$\boldsymbol{b}^{(i)} = \boldsymbol{a}^{(i)}\text{(所有 } i \neq r\text{),而 } \boldsymbol{b}^{(r)} = \boldsymbol{a}^{(r)} + \sum_{s \neq r} \lambda_s \boldsymbol{a}^{(s)},$$

则其行列式相等.

证 显然有

$$|B| = \sum \pm a_{q_1 1} \cdots \left(a_{q_r r} + \sum_{s \neq r} \lambda_s a_{q_r s} \right) \cdots a_{q_n n}$$
$$= |A| + \sum_{s \neq r} \lambda_s |C^{(s)}| = |A|,$$

因为这里 $C^{(s)}=[\boldsymbol{a}',\cdots,\boldsymbol{a}^{(r)}=\boldsymbol{a}^{(s)},\cdots,\boldsymbol{a}^{(n)}]$中 r 列与 s 列相同,从而有 $|C^{(s)}|=0$,得证.

注 对于行矢量的情况有类似结果.

(8) 若 $A=[\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)}]$中 $\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)}$ 是线性相关的,则其行列式为零.

证 若 $\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)}$ 线性相关,则根据定义必有一个矢量(比如说)$\boldsymbol{a}^{(r)}$ 可表达为其余矢量之线性组合,即

$$\boldsymbol{a}^{(r)} = \sum_{s \neq r} \lambda_s \boldsymbol{a}^{(s)},$$

于是由性质(7)可知

$$|A| = \sum_{s \neq r} \lambda_s |C^{(s)}| = 0.$$

3.3 行列式的展开

3.3.1 子行列式

设有 n 阶行列式 $A=\{a_{ij}\}$，和下列两个任意排列：

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ r_1 & r_2 & \cdots & r_n \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & \cdots & n \\ s_1 & s_2 & \cdots & s_n \end{pmatrix},$$

可以给出下面关于行列式的子式，主子式，余子式以及代数余子式诸概念的定义.

定义 1 从 n 阶行列式($|a_{ij}|, i, j=1,\cdots,n$)中相互无关地删去任何 $n-m$ 行(比如 $r_{m+1}<\cdots<r_n$ 诸行)和任何 $n-m$ 列(比如 $s_{m+1}<\cdots<s_n$ 诸列)($1\leqslant m\leqslant n$)，则留下的 m^2 个元素($a_{ij}, i=r_1,\cdots,r_m, j=s_1,\cdots,s_m$)所构成的 m 阶行列式，称为原行列式的 m 阶**子式**，记作 $M_{r_1\cdots r_m, s_1\cdots s_m}$.

定义 2 若从 n 阶行列式中删去任何 $n-m$ 行(比如 $r_{m+1}<\cdots<r_n$ 诸行)和对应的 $n-m$ 列(同为 $r_{m+1}<\cdots<r_n$ 的诸列)($m<n$)，则留下的 m^2 个元素($a_{ij}, i, j=r_1,\cdots,r_m$)所构成的 m 阶行列式，称为原行列式的 m 阶**主子式**，记作 $M_{r_1\cdots r_m}$.

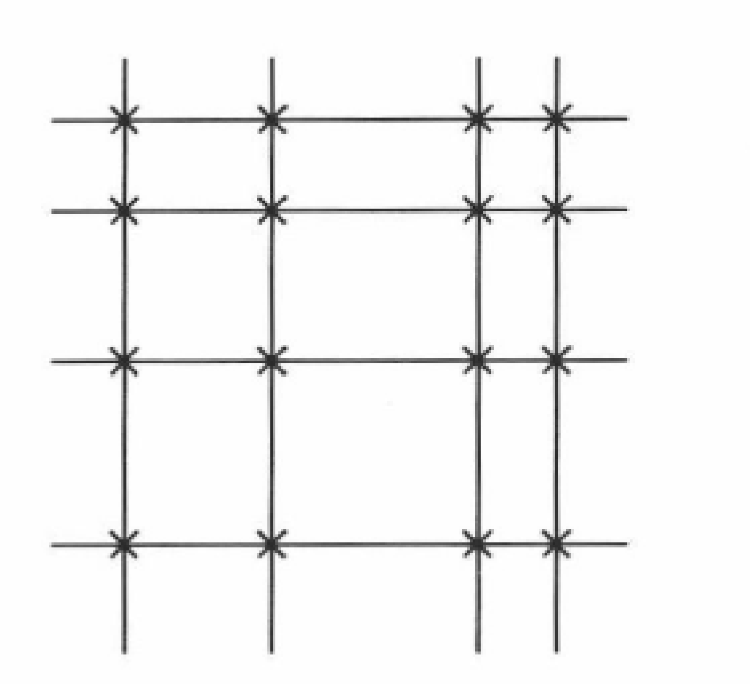
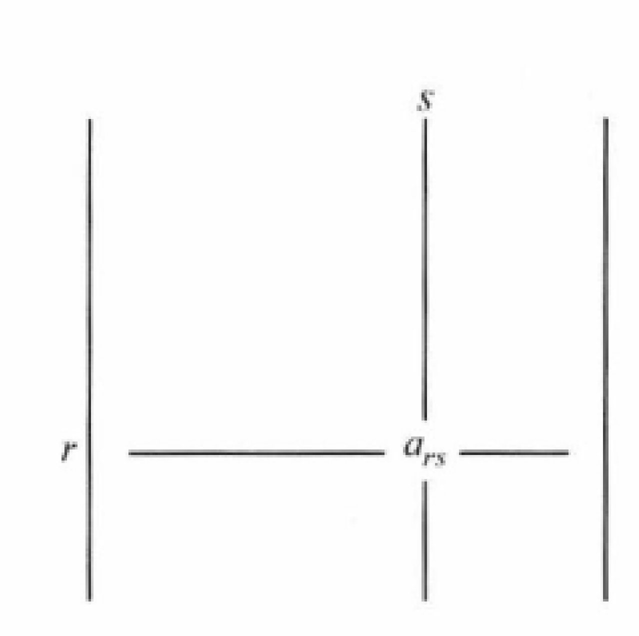
n 阶行列式总共有 $\binom{n}{m}^2=\left[\frac{n!}{m!\ (n-m)!}\right]^2$ 个 m 阶子式，其中有 $\binom{n}{m}=\frac{n!}{m!\ (n-m)!}$ 个是 m 阶主子式.

定义 3 如果 $n-m$ 阶子式中所保留的行和列($a_{ij}, i=r_{m+1},\cdots,r_n, j=s_{m+1},\cdots,s_n$)正好是获得 m 阶子式($a_{ij}=r_1,\cdots,r_m, j=s_1,\cdots,s_m$)时从 n 阶行列式($|a_{ij}|, i, j=1,\cdots,n$)中所删去的那些行和列，则称它们互为**余子式**，分别记作 M 和 M^c.

如下页左图中交叉点上元素所成行列式与非线上元素所成行列式即互为余子式.[注意到，主子式的余子式也是主子式.]

定义 4 设 n 阶行列式的一个 m 阶子式为 $M(|a_{ij}|, i=r_1,\cdots,r_m, j=s_1,\cdots,s_m)$，$M$ 的余子式为 $M^c(|a_{ij}|, i=r_{m+1},\cdots,r_n, j=s_{m+1},\cdots,s_n)$，则$(-)^{\sum_{\mu=1}^{m}(r_\mu+s_\mu)}M^c$ 称为 M 的**代数余子式**，记作 M^a 或 $\overline{M}$.[注意到，这里上标 $\sum_{\mu=1}^{m}(r_\mu+s_\mu)$ 等于 M 的所有对角元下标之和. 所以，对于主子式，其余子式亦为其代数余子式.]

特别是，相对于 a_{rs}(1 阶子式)的余子式是从 $|a_{ij}|$ 中删去 r 行和 s 列后所得行

列式(如上右图),而再乘以$(-)^{r+s}$后则得其代数余子式,后者通常记作 A_{rs}.

例如,5 阶行列式 $\det A=|a_{ij}|(i,j=1,\cdots,5)$,它有$\binom{5}{2}^2=100$ 个 2 阶子式,其中有 10 个是主子式,比如

$$M_{13}=\begin{vmatrix}a_{11} & a_{13}\\ a_{31} & a_{33}\end{vmatrix},\quad M_{13,25}=\begin{vmatrix}a_{12} & a_{15}\\ a_{32} & a_{35}\end{vmatrix},\text{等等};$$

3 阶子式也是 100 个,主子式也是 10 个,比如

$$M_{245}=\begin{vmatrix}a_{22} & a_{24} & a_{25}\\ a_{42} & a_{44} & a_{45}\\ a_{52} & a_{54} & a_{55}\end{vmatrix},\quad M_{245,134}=\begin{vmatrix}a_{21} & a_{23} & a_{24}\\ a_{41} & a_{43} & a_{44}\\ a_{51} & a_{53} & a_{54}\end{vmatrix},\text{等等};$$

还有$\binom{5}{1}^2=25$ 个 4 阶子式,其中有 5 个是主子式,比如,

$$M_{1245},\quad M_{1345,1234},\quad \text{等等}.$$

可以看出,其中 M_{13},M_{245},M_{1245}等是主子式;M_{13}与 M_{245},$M_{13,25}$与 $M_{245,134}$等分别互为余子式,即 $M_{245}=M_{13}^{c}$,$M_{245,134}=M_{13,25}^{c}$.

另外,M_{13}的代数余子式即为 M_{13}^{c},而 $M_{13,25}$的代数余子式是

$$M_{13,25}^{a}=\overline{M}_{13,25}=(-)^{(1+3)+(2+5)}M_{245,134}=-M_{245,134};$$

而 a_{25}的代数余子式是

$$A_{25}=(-)^{(2+5)}M_{1345,1234}=-M_{1345,1234}.$$

3.3.2 行列式按行(或列)展开

定理 设 $A=\{a_{rs}\}$为 n 阶矩阵,则

$$\det A=\sum_{s=1}^{n}a_{rs}A_{rs}\quad(1\leqslant r\leqslant n),$$

或类似地,

$$\det A=\sum_{r=1}^{n}a_{rs}A_{rs}\quad(1\leqslant s\leqslant n),$$

其中 A_{rs} 是 a_{rs} 的代数余子式.

证明 首先注意到行列式的定义(Ⅰ):

$$\det A=\sum \operatorname{sgn}\begin{pmatrix}1 & \cdots & n\\ p_1 & \cdots & p_n\end{pmatrix}a_{1p_1}\cdots a_{np_n},$$

其中每项均含一个 r 行元素.因此,将所有含 a_{rs} 的项集合在一起并暂记为 $a_{rs}\hat{A}_{rs}$;若能证明对一切 r,s 均有 $\hat{A}_{rs}=A_{rs}$,则定理即得证.

从证明 $\hat{A}_{11}=A_{11}$ 开始.从定义式可见,

$$a_{11}\hat{A}_{11}=a_{11}\sum_{(p_2\cdots p_n)}\operatorname{sgn}\begin{pmatrix}1 & 2 & \cdots & n\\ 1 & p_2 & \cdots & p_n\end{pmatrix}a_{2p_2}\cdots a_{np_n},$$

其中 $\sum\limits_{(p_2\cdots p_n)}$ 为对一切可能的置换 $\begin{pmatrix}2 & \cdots & n\\ p_2 & \cdots & p_n\end{pmatrix}$ 求和,而 $\operatorname{sgn}\begin{pmatrix}1 & 2 & \cdots & n\\ 1 & p_2 & \cdots & p_n\end{pmatrix}=\operatorname{sgn}\begin{pmatrix}2 & \cdots & n\\ p_2 & \cdots & p_n\end{pmatrix}$,可见 $\sum$ 部分正好是删去第一行和第一列后所得行列式,即 a_{11} 的余子式 M_{11}^{c},因 $(-)^{1+1}=+$,所以的确有 $\hat{A}_{11}=A_{11}$.

至于一般情况,只要通过行轮换 $(1\cdots r)$(或连续 $r-1$ 次行对换)与列轮换 $(1\cdots s)$(或连续 $s-1$ 次列对换)将 a_{rs} 移至 $(1,1)$ 位置,由于其余的行和列之相对位置均未变,即 a_{rs} 的余子式 M_{rs}^{c} 未变,而置换(轮换或对换)提供一因子 $(-)^{(r-1)+(s-1)}$,利用上一段结果得

$$\hat{A}_{rs}=(-)^{r+s}M_{rs}^{c}=A_{rs}.$$

至此,按列(或按行)相加,则定理全部得证.

推论 1 $\det A(\boldsymbol{a}^{(s)}\to\boldsymbol{x})=\sum\limits_{r}x_rA_{rs}$,

$\det A(\widetilde{\boldsymbol{a}}_{(r)}\to\widetilde{\boldsymbol{x}})=\sum\limits_{s}x_sA_{rs}$.

这里 $\boldsymbol{a}^{(s)}\to\boldsymbol{x}$ 指 A 中 s 列元素 a_{rs} 用 x_r 代替,

$\widetilde{\boldsymbol{a}}_{(r)}\to\widetilde{\boldsymbol{x}}$ 指 A 中 r 行元素 a_{rs} 用 x_s 代替.

推论 2 $\sum\limits_{s}a_{ts}A_{rs}=0\quad(t\neq r)$,

$\sum\limits_{r}a_{rt}A_{rs}=0\quad(t\neq s)$.

因为有 t,r 二行(前者)或 t,s 二列(后者)相同.

定理与推论 2 结合起来可写成

$$\sum_{s}a_{ts}A_{rs}=\delta_{tr}(\det A),$$

$$\sum_r a_{tr} A_{rs} = \delta_{ts}(\det A),$$

这里 δ_{ij} 是 Kronecker δ **符号**

$$\delta_{ij} = \begin{cases} 1 & 若\ i = j, \\ 0 & 若\ i \neq j. \end{cases}$$

作为按行展开定理之应用的一个特例，考虑下列 n 阶行列式

$$\det D = \begin{vmatrix} d_{11} & 0 & \cdots & 0 \\ d_{21} & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{vmatrix},$$

其对角线右上角元素全为 0，称为**下三角阵行列式**. 不断应用按行展开定理，可得

$$\det D = d_{11} d_{22} \cdots d_{nn}.$$

只有对角元素不为零的矩阵称为**对角矩阵**，记作 $\mathrm{diag}[d_{11}, d_{22}, \cdots, d_{nn}]$，于是立即得到

$$\det \mathrm{diag}[d_{11}, d_{22}, \cdots, d_{nn}] = d_{11} d_{22} \cdots d_{nn}.$$

3.3.3 行列式的 Laplace 展开①

令 m 阶子式 M 具有行标 $r_1 < r_2 < \cdots < r_m$ 和列标 $s_1 < s_2 < \cdots < s_m$，其 $n-m$ 阶余子式 M^c 具有行标 $r_{m+1} < \cdots < r_n$ 和列标 $s_{m+1} < \cdots < s_n$. 若 $\sigma_1, \cdots, \sigma_m$ 为 $s_1, \cdots, s_m$ 上的一个置换，$\sigma_{m+1}, \cdots, \sigma_n$ 为 $s_{m+1}, \cdots, s_n$ 上的一个置换；则有

$$M = \sum_{(\sigma_1, \cdots, \sigma_m)} \mathrm{sgn} \begin{pmatrix} s_1 & \cdots & s_m \\ \sigma_1 & \cdots & \sigma_m \end{pmatrix} a_{r_1 \sigma_1} \cdots a_{r_m \sigma_m},$$

$$M^c = \sum_{(\sigma_{m+1}, \cdots, \sigma_n)} \mathrm{sgn} \begin{pmatrix} s_{m+1} & \cdots & s_n \\ \sigma_{m+1} & \cdots & \sigma_n \end{pmatrix} a_{r_{m+1} \sigma_{m+1}} \cdots a_{r_n \sigma_n}.$$

这里求和号 $\sum$ 包括相应的全部可能置换. 乘积 MM^c 包括 $a_{r_1\sigma_1} \cdots a_{r_m\sigma_m} a_{r_{m+1}\sigma_{m+1}} \cdots a_{r_n\sigma_n}$ 类型的项，具有符号

$$\mathrm{sgn} = \mathrm{sgn} \begin{pmatrix} s_1 & \cdots & s_m \\ \sigma_1 & \cdots & \sigma_m \end{pmatrix} \mathrm{sgn} \begin{pmatrix} s_{m+1} & \cdots & s_n \\ \sigma_{m+1} & \cdots & \sigma_n \end{pmatrix}.$$

因此，在 $\mathrm{sgn} \begin{pmatrix} r_1 & \cdots & r_n \\ s_1 & \cdots & s_n \end{pmatrix} MM^c$ 中，$a_{r_1\sigma_1} \cdots a_{r_n\sigma_n}$ 这项具有所指定的符号. 于是，行列式 $|A|$ 可展开为

$$|A| = \sum \mathrm{sgn} \begin{pmatrix} r_1 & \cdots & r_n \\ s_1 & \cdots & s_n \end{pmatrix} MM^c,$$

① G. Kowalewski, *Determinant Theory*, p. 32.

这里，$\sum$ 是在行指标 $r_1\cdots r_n$ 的给定划分下对 $s_1\cdots s_m$ 的 $\binom{n}{m}$ 个可能的划分求和.

可以按照与上一小节类似的方式证明 $\operatorname{sgn}\begin{pmatrix} r_1 & \cdots & r_n \\ s_1 & \cdots & s_n \end{pmatrix} M^c$ 是 M 的代数余子式 M^a. 首先，若 M 为主子式，则 M^c 也是主子式，因而有 $\operatorname{sgn}\begin{pmatrix} r_1 & \cdots & r_n \\ r_1 & \cdots & r_n \end{pmatrix} = +$，$M^c = M^a$；即此时上述结果是正确的. 其次，若 M 非主子式，可从 r_1 行开始，通过与其前诸行实施行对换，依次进行下去，直至 $r_1\cdots r_m$ 位于前 m 行，共需 $\sum_{\mu=1}^{m}(r_\mu-\mu)$ 次行对换；再类似地实施列对换，使之位于前 m 列，共需 $\sum_{\mu=1}^{m}(s_\mu-\mu)$ 次列对换；同时注意到，这种步骤保持 M 和 M^c 中元素的相对位置不变，而仅将其变为主子式. 结果得到

$$\operatorname{sgn}\begin{pmatrix} r_1 & \cdots & r_n \\ s_1 & \cdots & s_n \end{pmatrix} M^c = (-)^{\sum_{\mu=1}^{m}(r_\mu+s_\mu)} M^c = M^a.$$

最后得到下述 **Laplace 展开定理**，亦称 **Laplace 展开公式**.

Laplace 展开定理　一个 n 阶行列式等于任选 m 行(或列)下($1\leqslant m\leqslant n$)一切可能的$\left[共\binom{n}{m}个\right]$$m$ 阶子式与其代数余子式乘积之和，即

$$\det A = \sum MM^a.$$

应该注意到，上一小节关于行列式按行(或列)展开的定理是 Laplace 定理的一个特例($m=1$).

作为 Laplace 展开定理之应用的另一特例，考虑下列 $n+m$ 阶行列式：

$$\det D = \begin{vmatrix} a_{11} & \cdots & a_{1n} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1n} & b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} & b_{m1} & \cdots & b_{mm} \end{vmatrix} = \begin{vmatrix} A & 0 \\ C & B \end{vmatrix},$$

右边是缩写，其中 $A=\{a_{ij}\}$，$B=\{b_{ij}\}$，$C=\{c_{ij}\}$分别为 n 阶、m 阶、$m\times n$ 阶矩阵，而 0 则为 $n\times m$ 阶 0 矩阵；所以右边形式称为**下三角块阵**的行列式. 应用 Laplace 展开定理，对其前 n 行展开，并注意到主子式的代数余子式即其余子式，立即得到此行列式为两对角块方阵的行列式之积，

$$\det D = (\det A)\cdot(\det B);$$

更一般地，主对角块为方阵的下三角块阵之行列式为其对角块行列式之积.

关于一般分块矩阵的运算见 3.4 节.

3.3.4 行列式值的计算——凝聚法

设 n 阶行列式 $|A|=\det\{a_{ik}\}$，若其首元不为零 $a_{11}\neq 0$，则可利用行列式的性质，按下列步骤将之降阶至 $n-1$ 阶行列式：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}\begin{vmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \frac{a_{21}}{a_{11}} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{a_{11}} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad \text{性质(4)}$$

$$= a_{11}\begin{vmatrix} 1 & 0 & \cdots & 0 \\ \frac{a_{21}}{a_{11}} & a_{22}-\frac{a_{21}}{a_{11}}a_{12} & \cdots & a_{2n}-\frac{a_{21}}{a_{11}}a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{a_{11}} & a_{n2}-\frac{a_{n1}}{a_{11}}a_{12} & \cdots & a_{nn}-\frac{a_{n1}}{a_{11}}a_{1n} \end{vmatrix} \quad \text{性质(7)}$$

$$= a_{11}\begin{vmatrix} a_{22}-\frac{a_{21}}{a_{11}}a_{12} & \cdots & a_{2n}-\frac{a_{21}}{a_{11}}a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n2}-\frac{a_{n1}}{a_{11}}a_{12} & \cdots & a_{nn}-\frac{a_{n1}}{a_{11}}a_{1n} \end{vmatrix} \quad \text{按行展开定理（3.3.2 节）}$$

$$= \frac{a_{11}}{a_{11}^{n-1}}\begin{vmatrix} a_{11}a_{22}-a_{12}a_{21} & \cdots & a_{11}a_{2n}-a_{1n}a_{21} \\ \vdots & \ddots & \vdots \\ a_{11}a_{n2}-a_{12}a_{n1} & \cdots & a_{11}a_{nn}-a_{1n}a_{n1} \end{vmatrix} \quad \text{性质(4)}$$

$$= \frac{1}{a_{11}^{n-2}}\begin{vmatrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} & \cdots & \begin{vmatrix} a_{11} & a_{1n} \\ a_{21} & a_{2n} \end{vmatrix} \\ \vdots & \ddots & \vdots \\ \begin{vmatrix} a_{11} & a_{12} \\ a_{n1} & a_{n2} \end{vmatrix} & \cdots & \begin{vmatrix} a_{11} & a_{1n} \\ a_{n1} & a_{nn} \end{vmatrix} \end{vmatrix}$$

$$= \frac{1}{a_{11}^{n-2}}\det\left\{\begin{vmatrix} a_{11} & a_{1k} \\ a_{i1} & a_{ik} \end{vmatrix}\right\}.$$

因而，n 阶行列式 $|A|$ 降至 $n-1$ 阶行列式 $|\overset{①}{A}|$，即

$$\overset{n\times n}{|A|}=a_{11}^{2-n}\ \overset{\overset{n-1\times n-1}{①}}{|A|},\quad \overset{①}{a}_{ik}=\begin{vmatrix}a_{11} & a_{1k}\\ a_{i1} & a_{ik}\end{vmatrix},$$

其中$|\overset{①}{A}|=\det\{\overset{①}{a}_{ik}\}$，而$\overset{①}{a}_{ik}(i,k=2,\cdots,n)$是通过相对于首元$a_{11}$凝聚而得.

按此步骤继续进行下去，$n-1$次后得到

$$\overset{n\times n}{|A|}=a_{11}^{2-n}\ \overset{①}{a}_{22}^{3-n}\ \overset{②}{a}_{33}^{4-n}\cdots\overset{(n-1)}{a}_{nn},$$

注意到，$|\overset{①}{A}|$的首元为$\overset{①}{a}_{22}$，$|\overset{②}{A}|$的首元为$\overset{②}{a}_{33}$，而

$$\overset{②}{a}_{ik}=\begin{vmatrix}\overset{①}{a}_{22} & \overset{①}{a}_{2k}\\ \overset{①}{a}_{i2} & \overset{①}{a}_{ik}\end{vmatrix}\quad(i,k=3,\cdots,n);$$

依此类推.

另外，如果遇到$a_{11}=0$(或$\overset{①}{a}_{22}=0$等)的情况，若有某个元素$a_{ik}\neq 0$，则可通过将i行与1行对换$(i\leftrightarrow 1)$和k列与1列对换$(k\leftrightarrow 1)$而将a_{ik}移至首位，即可按凝聚法进行计值.

凝聚法尤其适宜于计算机上对行列式进行计算. 计算过程中的每一步，当然会首先将行列式元素$\overset{①}{a}_{ik}\cdots$等算出，再继续进行下去.

备注 实际数值计算中，若每一步均选主元(元素中数值与准确度最大者)作为首元，会使计算准确度大为提高. 具体情况可参考下章4.2.1节备注及习题4.2.

3.4 矩阵的分块运算

可以将一个矩阵划分成若干块，每个分块称为一个**子矩阵**.

例如，对于一个$m\times n$矩阵$A=\{a_{ij}\}$，若将m和n按顺序划分为

$$m=m_1+m_2+\cdots+m_r,\quad n=n_1+n_2+\cdots+n_s,$$

则可将A写成由$r\times s$个子矩阵$\mathscr{A}_{\rho\sigma}$($\mathscr{A}_{\rho\sigma}$为$m_\rho\times n_\sigma$矩阵)作为块元素而构成的$r\times s$**分块矩阵**$A=\{\mathscr{A}_{\rho\sigma}\}$. 类似地，对于一个$n\times l$矩阵$B=\{b_{jk}\}$，若

$$n=n_1+n_2+\cdots+n_s,\quad l=l_1+l_2+\cdots+l_t,$$

则B也可写为$s\times t$分块矩阵$B=\{\mathscr{B}_{\sigma\tau}\}$，其中$\mathscr{B}_{\sigma\tau}$是$n_\sigma\times l_\tau$矩阵.

3.4.1 矩阵的分块乘法

对于分块矩阵可以像普通矩阵那样进行运算. 特别是对于上面所述矩阵A和B，有下列**分块链式积**(关键是A的列与B的行要作同样划分)：

$$AB=\left\{\sum_{\sigma=1}^{s}\mathscr{A}_{\rho\sigma}\mathscr{B}_{\sigma\tau}\right\},$$

它是 $r\times t$ 分块矩阵. 对块元素再实施链乘并取其 c_{ik} 元素，并注意取 i 行所在的 $\mathscr{A}$ 块，k 列所在的 $\mathscr{B}$ 块，于是

$$c_{ik}=\Big(\sum_{\sigma=1}^{s}\mathscr{A}_{\rho\sigma}\mathscr{B}_{\sigma\tau}\Big)_{ik}=\sum_{j=1}^{n_1}a_{ij}b_{jk}+\sum_{j=n_1+1}^{n_1+n_2}a_{ij}b_{jk}+\cdots+\sum_{j=\sum_{\nu=1}^{s-1}n_\sigma+1}^{n}a_{ij}b_{jk}$$

$$=\sum_{j=1}^{n}a_{ij}b_{jk}=(AB)_{ik},$$

这证明链乘的上述分块运算是正确的.

3.4.2 同阶矩阵之积的行列式

现在来讨论矩阵之积的行列式，为简单起见，只考虑同阶矩阵的情形；有下列定理.

定理 同阶矩阵之积的行列式等于各矩阵的行列式之积.

证 设 A,B,C 均为 n 阶矩阵，欲证明：

$$若\quad C=AB,\quad 则\quad |C|=|A||B|.$$

作下列 $2n$ 阶行列式 $|\Gamma|$：

$$|\Gamma|=\begin{vmatrix}A & 0\\ -I_n & B\end{vmatrix},$$

其中 I_n 是 n 阶单位矩阵，0 为 n 阶零矩阵. 根据 Laplace 展开定理得

$$|\Gamma|=|A||B|.$$

另一方面，利用行列式的性质(7)，重复用 b_{jk} 乘 $|\Gamma|$ 的第 j 列($j=1,2,\cdots,n$)，累计加至 $|\Gamma|$ 的第($n+k$)列，再对 $k=1,2,\cdots,n$ 进行重复运算，以使行列式变为

$$|\Gamma|=\begin{vmatrix}A & AB\\ -I_n & 0\end{vmatrix}=\begin{vmatrix}A & C\\ -I_n & 0\end{vmatrix}.$$

再利用行列式的性质(2)，二行(或二列)互换改变符号，重复将第 k 列与第($n+k$)列互换($k=1,2,\cdots,n$)得到

$$|\Gamma|=(-1)^n\begin{vmatrix}C & A\\ 0 & -I_n\end{vmatrix}=(-1)^n\,|C||-I_n|$$

$$=(-)^{2n}\,|C|=|C|;$$

这里除利用了 Laplace 展开定理外，还利用了行列式的性质(4)，即 $|\lambda A|=\lambda^n|A|$.

根据以上各式，即得欲证明的结果.

另外，显然可见，此结果可以推广到若干个同阶矩阵乘积的情况.

同时，由于 $A^{-1}A=I$，故有($\det A^{-1}$)=($\det A$)$^{-1}$.

3.4.3 同阶行列式的乘积

前面的定理：若 $C=AB$，则 $|C|=|A||B|$，可以看作两行列式 $|A|$ 与 $|B|$ 乘积 $|A||B|$ 的乘法定则：利用 $C=AB$ 行乘列法计算

$$c_{ik}=\sum_j a_{ij}b_{jk},$$

再取其行列式. 也就是说，

$$|a_{ij}||b_{jk}|=\left|\sum_j a_{ij}b_{jk}\right|.$$

显然，此结果可以推广到若干个同阶行列式相乘的情况.

3.4.4 分块矩阵的行列式

令人特别感兴趣的是所有对角块均为方阵的分块矩阵，这一小节讨论分块矩阵的行列式.

设 A 为 n 阶矩阵，划分为 $n=p+q$ 的 2×2 分块矩阵，即 $A=\{a_{ij}\}=\{\mathscr{A}_{\rho\sigma}\}$，其中 $\mathscr{A}_{11}$，$\mathscr{A}_{22}$ 为方阵. 若 $\mathscr{A}_{11}$ 为非奇异矩阵，则有

$$\begin{aligned}\det A&=\det\begin{bmatrix}\mathscr{A}_{11} & \mathscr{A}_{12}\\ \mathscr{A}_{21} & \mathscr{A}_{22}\end{bmatrix}\\ &=\det\mathscr{A}_{11}\det(\mathscr{A}_{22}-\mathscr{A}_{21}\mathscr{A}_{11}^{-1}\mathscr{A}_{12}).\end{aligned}$$

证明 通过矩阵的分块乘法容易验证下列关系：

$$\begin{bmatrix}I & 0\\ \mathscr{A}_{21}\mathscr{A}_{11}^{-1} & I\end{bmatrix}\begin{bmatrix}\mathscr{A}_{11} & 0\\ 0 & \mathscr{A}_{22}-\mathscr{A}_{21}\mathscr{A}_{11}^{-1}\mathscr{A}_{12}\end{bmatrix}\begin{bmatrix}I & \mathscr{A}_{11}^{-1}\mathscr{A}_{12}\\ 0 & I\end{bmatrix}=\begin{bmatrix}\mathscr{A}_{11} & \mathscr{A}_{12}\\ \mathscr{A}_{21} & \mathscr{A}_{22}\end{bmatrix}.$$

利用上述关系并注意到块三角形矩阵等于其对角块之积，以及应用 3.4.2 节矩阵之积的行列式的定理，即得证.

若 $\mathscr{A}_{22}$ 为非奇异矩阵，类似地有

$$\det\begin{vmatrix}\mathscr{A}_{11} & \mathscr{A}_{12}\\ \mathscr{A}_{21} & \mathscr{A}_{22}\end{vmatrix}=\det\mathscr{A}_{22}\det(\mathscr{A}_{11}-\mathscr{A}_{12}\mathscr{A}_{22}^{-1}\mathscr{A}_{21}).$$

3.5 矩 阵 的 秩

3.5.1 秩的定义

定义 对于 $m\times n$ 矩阵 A，若其全体 p 阶子式均为零，但是至少有一个 r 阶子式不为零（$r<p\leqslant\min(m,n)$），则称此矩阵的**秩**为 r，记作 r_A，$r(A)$ 或 $\operatorname{rank}A$. 另外，若 $r=\min(m,n)$，则称此矩阵具有**满秩**.

简而言之，矩阵的秩等于其最大不为零子式的阶.

注意，这里 p 阶子式的概念指的是从 $m\times n$ 矩阵 A 中删去任何 $m-p$ 行与任何 $n-p$ 列后所留下的 p^2 个元素构成的行列式；是行列式中子式概念的推广.

还应指出，零矩阵 0 是一特例，其秩为 0 而小于 1，因为其 1 阶子式均为零（即 $0_{ik}=0$）.

3.5.2 满秩方阵的有关定理

对于满秩方阵 A，其行列式 $|A|\neq 0$，这时 $A=\{\boldsymbol{a}',\boldsymbol{a}'',\cdots,\boldsymbol{a}^{(n)}\}$ 的 n 个列矢量 $\boldsymbol{a}',\cdots,\boldsymbol{a}^{(n)}$ 是线性无关的；也就是说，A 是非奇异的. 因为，若 $\boldsymbol{a}',\cdots,\boldsymbol{a}^{(n)}$ 线性相关，则根据 3.2 节行列式的性质(8)知，必有 $|A|=0$，与满秩的前提矛盾. 另一方面，令 $A=\begin{bmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\\ \vdots\\ \boldsymbol{b}_n\end{bmatrix}$，则行矢量 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n$ 也是线性无关的. 这是因为 $A^{\mathrm{T}}=\{\boldsymbol{b}_1^{\mathrm{T}},\cdots,\boldsymbol{b}_n^{\mathrm{T}}\}$，若 $\boldsymbol{b}_1^{\mathrm{T}},\cdots,\boldsymbol{b}_n^{\mathrm{T}}$ 线性相关，就有 $\lambda_1\boldsymbol{b}_1^{\mathrm{T}}+\cdots+\lambda_n\boldsymbol{b}_n^{\mathrm{T}}=0$，其中 $\lambda_1,\cdots,\lambda_n$ 不全为零，则 $|\boldsymbol{A}^{\mathrm{T}}|=0$；利用 $|A|=|A^{\mathrm{T}}|$，导致 $|A|=0$，与满秩的前提矛盾. 于是有关于满秩方阵的下述定理.

定理 满秩方阵（$|A|\neq 0, r=n$）是非奇异矩阵，其全体列矢量（行矢量）是线性无关的.

3.5.3 列秩与行秩及有关定理①

定义 一个 $m\times n$ 矩阵 A，其 n 个列矢量（m 个行矢量）中线性无关矢量的个数 $r_{\mathrm{c}}(r_{\mathrm{r}})$ 称为**列秩（行秩）**；或记为 $r_{\mathrm{c}}(A)(r_{\mathrm{r}}(A))$.

关于矩阵的列秩和行秩有下述定理.

定理 矩阵的行秩与列秩相等，而且等于矩阵的秩.

证 设 $m\times n$ 矩阵 $A=\{\boldsymbol{a}^{(1)},\cdots,\boldsymbol{a}^{(n)}\}$，其秩为 r 而列秩为 r_{c}、行秩为 r_{r}. 不失一般性，令 $\boldsymbol{a}^{(1)},\cdots,\boldsymbol{a}^{(r_{\mathrm{c}})}$ 为线性无关列矢量，并令 $m\times r_{\mathrm{c}}$ 矩阵 $C=\{\boldsymbol{a}^1,\cdots,\boldsymbol{a}^{(r_{\mathrm{c}})}\}$. 因而 $\boldsymbol{a}^{(k)}$ 可表达为 $\boldsymbol{a}^{(1)},\cdots,\boldsymbol{a}^{(r_{\mathrm{c}})}$ 的线性组合，可写成

$$\boldsymbol{a}^{(k)}=\sum_{l=1}^{r_{\mathrm{c}}}\boldsymbol{a}^{(l)}\lambda_l^{(k)}=C\boldsymbol{\lambda}^{(k)}\quad(k=1,\cdots,n);$$

① 参考：A. S. Deif, *Advanced Matrix Theory for Scientists and Engineers*, Abacus Press, London, 1982, p.47.

其中 $\boldsymbol{\lambda}^{(k)}$ 为 r_c 维列矢量，$\boldsymbol{\lambda}^{(k)}=\begin{Bmatrix}\lambda_1^{(k)}\\ \vdots \\ \lambda_{r_c}^{(k)}\end{Bmatrix}$. 再令 $r_c\times n$ 矩阵 $\Lambda=\{\boldsymbol{\lambda}^{(1)},\cdots,\boldsymbol{\lambda}^{(n)}\}$，则矩阵 A 可表达为

$$A=\{\boldsymbol{a}^{(1)},\cdots,\boldsymbol{a}^{(n)}\}=\{C\boldsymbol{\lambda}^{(1)},\cdots,C\boldsymbol{\lambda}^{(n)}\}=C\{\boldsymbol{\lambda}^{(1)},\cdots,\boldsymbol{\lambda}^{(n)}\}=C\Lambda.$$

这意味着 A 的行矢量是 Λ 的行矢量之线性组合. 由此可见，A 的行秩不能超过 Λ 中行矢量之个数，即 $r_r\leqslant r_c$. 考虑 A 的转置 A^{T}，类似地得到 $r_c\leqslant r_r$. 因而有

$$r_r=r_c,$$

即矩阵的行秩与列秩相等，定理之前段得证.

再来证定理之后段. 同样，不失一般性，可再令 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_{r_r}$ 为 A 中的线性无关行矢量. 由于已证明 $r_r=r_c$，显然，下列 r_c 阶行列式

$$\begin{vmatrix}a_{11} & \cdots & a_{1r_c}\\ \vdots & \ddots & \vdots\\ a_{r_c1} & \cdots & a_{r_cr_c}\end{vmatrix}\neq 0,$$

而任何 $p(>r_c)$阶行列式均为零，故 A 之秩 $r=r_c$.

最后结果是

$$r_r=r_c=r.$$

定理全部得证.

3.6 矩阵求逆

对于非奇异 n 阶矩阵 A，有唯一的逆矩阵 A^{-1}使

$$AA^{-1}=A^{-1}A=I.$$

下面介绍计算矩阵之逆的具体方法.

3.6.1 利用伴随矩阵求逆

由 n 阶矩阵 $A=\{a_{ij}\}$中元素 a_{ij} 的代数余子式 A_{ij} 为元素构成的下列矩阵

$$\begin{bmatrix}A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \vdots & \vdots & \ddots & \vdots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{bmatrix},$$

称为矩阵 A 的**伴随矩阵**，记作 $\mathrm{adj}A$，即（T 或～表示转置）

$$\mathrm{adj}A=\{A_{ij}\}^{\mathrm{T}}=\{\widetilde{A}_{ij}\}.$$

由 3.3.2 节推论 2 已经知道，

$$\sum_s a_{ts} A_{rs} = \delta_{tr}(\det A),$$

而左边可写成

$$\sum_s a_{ts} \widetilde{A}_{sr} = (A\ \mathrm{adj}A)_{tr},$$

所以

$$A(\mathrm{adj}A) = (\det A)I;$$

即

$$A^{-1} = \frac{\mathrm{adj}A}{\det A} \quad 或 \quad (A^{-1})_{rs} = \frac{A_{sr}}{\det A}.$$

由此可知，若 $\det A \neq 0$，则 A 是非奇异矩阵，若 $\det A = 0$，则 A 是奇异矩阵.

3.6.2 利用矩阵的变换求逆

若 A 为非奇异矩阵，则有非奇异矩阵 Q 和 P，使

$$QAP = I,$$
$$A = Q^{-1}P^{-1},$$
$$A^{-1} = PQ.$$

具体计算格式如下：

I	A
	I

Q	I
	P

从左边格式开始，对 A 施加行变换时对其左边 I 施加同一变换，而对 A 施加列变换时则对其下边 I 施加同一变换，经若干次行和列变换使 A 化为单位阵 I 时，其左边化为 Q，其下边化为 P，于是求得 $A^{-1} = PQ$.

例如，

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 3 \end{bmatrix},$$

依次经过下列变换：$R_1 \leftrightarrow R_2$，$C_2 - 2C_1$ 及 $C_3 - C_1$，$R_2 - 2R_1$ 及 $R_3 + R_1$，$\frac{1}{2}R_3$ 及 $R_2 \leftrightarrow R_3$，$C_3 - 2C_2$，$R_3 + 5R_2$，$\frac{1}{8}R_3$（R 和 C 分别代表行和列）后 A 化为 I，同时得到

$$Q = \frac{1}{16}\begin{bmatrix} 0 & 16 & 0 \\ 0 & 8 & 8 \\ 2 & 1 & 5 \end{bmatrix} \quad 和 \quad P = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix},$$

所以，

$$A^{-1}=PQ=\frac{1}{16}\begin{bmatrix}6&3&-1\\-4&6&-2\\2&1&5\end{bmatrix}.$$

3.6.3 利用矩阵的分块运算求逆

设 n 阶非奇异矩阵 A 划分为 2×2 分块矩阵

$$A=\begin{bmatrix}\mathscr{A}_{11}&\mathscr{A}_{12}\\\mathscr{A}_{21}&\mathscr{A}_{22}\end{bmatrix},$$

若 $\mathscr{A}_{11}$ 和 $\mathscr{A}_{22}$ 均为非奇异方阵，容易得到

$$A^{-1}=\begin{bmatrix}(\mathscr{A}_{11}-\mathscr{A}_{12}\mathscr{A}_{22}^{-1}\mathscr{A}_{21})^{-1}&-\mathscr{A}_{11}^{-1}\mathscr{A}_{12}(\mathscr{A}_{22}-\mathscr{A}_{21}\mathscr{A}_{11}^{-1}\mathscr{A}_{12})^{-1}\\-\mathscr{A}_{22}^{-1}\mathscr{A}_{21}(\mathscr{A}_{11}-\mathscr{A}_{12}\mathscr{A}_{22}^{-1}\mathscr{A}_{21})^{-1}&(\mathscr{A}_{22}-\mathscr{A}_{21}\mathscr{A}_{11}^{-1}\mathscr{A}_{12})^{-1}\end{bmatrix}.$$

例如，很容易求得

$$A^{-1}=\begin{bmatrix}1&0&6&8\\0&1&5&4\\0&0&-1&0\\0&0&0&-1\end{bmatrix}^{-1}=\begin{bmatrix}I_2&\begin{bmatrix}6&8\\5&4\end{bmatrix}\\0&-I_2\end{bmatrix}^{-1}$$

$$=\begin{bmatrix}I_2&\begin{bmatrix}6&8\\5&4\end{bmatrix}\\0&-I_2\end{bmatrix}=A;$$

中间一步也可对分块矩阵利用伴随矩阵法求得.

3.6.4 逐步求近法

上述三种方法，在给定矩阵元为准确值和没有舍入误差的假设下，经过有限次运算就可得到准确解. 但在实际情况中，给定矩阵元往往是近似值，而且大量计算中的舍入误差不可避免，所得均为近似解；为了保持一定的准确度，需采用逐步求近法.

例如，对矩阵 A，求得一近似逆 B_0 后可定义一残差矩阵

$$R_0=I-AB_0,$$

其矩阵元很小，但仅当 B_0 为 A 的准确逆 A^{-1} 时才为零. 于是可构造以下序列：

$$B_k=B_{k-1}(I+R_{k-1}),\quad R_k=I-AB_k,\quad k=1,2,\cdots;$$

于是，容易证明

$$B_k=A^{-1}(I-R_0^{2^k}).$$

因为 R_0 的范数 $\|R_0\|<1$，若 k 充分大，则 B_k 将经由迭代而逼近 A^{-1}. [**注** 关于 $m\times$

n 矩阵 $R=\{r_{ij}\}$的 Hölder 范数有 $\|R\|_1=\max\limits_{1\leqslant j\leqslant n}\sum\limits_{i=1}^{m}|r_{ij}|$，$\|R\|_\infty=\max\limits_{1\leqslant i\leqslant m}\sum\limits_{j=1}^{n}|r_{ij}|$.]

3.7 矩阵的迹

方阵 A 的主对角线上诸元素之和称为**矩阵的迹**，记作 $\mathrm{tr}A$ 或 $\mathrm{Sp}A$（分别来自英文 trace 和德文 Spur），即

$$\mathrm{tr}A=\sum_i a_{ii}.$$

可以证明，有限个矩阵之积在矩阵的任何**轮换**下，其迹为不变量.

首先，设 $A=\{a_{ij}\}$和 $B=\{b_{ji}\}$分别为 $m\times n$ 和 $n\times m$ 矩阵，则有

$$\mathrm{tr}(AB)=\sum_{i=1}^{m}\Big(\sum_{j=1}^{n}a_{ij}b_{ji}\Big)=\sum_{j=1}^{n}\Big(\sum_{i=1}^{m}b_{ji}a_{ij}\Big)=\mathrm{tr}(BA).$$

其次，设 $A_1,A_2,\cdots,A_k$ 为 k 个矩阵，使积 $A_1A_2,A_2A_3,\cdots,A_kA_1$ 可定义，若令

$$C=A_1A_2\cdots A_r,\quad D=A_{r+1}\cdots A_k,\quad 1\leqslant r\leqslant k-1.$$

则由上式可得 $\mathrm{tr}(CD)=\mathrm{tr}(DC)$，代入即得

$$\mathrm{tr}(A_1A_2\cdots A_k)=\mathrm{tr}(A_{r+1}\cdots A_kA_1\cdots A_r),\quad 1\leqslant r\leqslant k-1.$$

上述结果得证. 注意到此结果仅在轮换下成立.

3.8 若干特种行列式

3.8.1 Vandermonde 行列式

下列行列式

$$\begin{vmatrix} 1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\\ \vdots & \vdots & \ddots & \vdots\\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1}\end{vmatrix}$$

称为 n 阶 **Vandermonde 行列式**，记作 V_n. 首先，应用 3.2 节行列式的性质(7)，将 i 行乘以 $-x_n$ 加至 $i+1$ 行$[i=1,2,\cdots,(n-1)]$，最后一列$(j=n)$变为$[1\ \ 0\ \ \cdots\ \ 0]^{\mathrm{T}}$. 按 n 列作展开并提出余子式中各列的公因子 $\prod\limits_{j=1}^{n-1}(x_j-x_n)$ 后得到

$$\mathrm{V}_n=(-)^{n+1}\prod_{j=1}^{n-1}(x_j-x_n)\cdot\mathrm{V}_{n-1}=\prod_{j=1}^{n-1}(x_n-x_j)\cdot\mathrm{V}_{n-1},$$

其中$(-)^{n+1}$是代数余子式的符号因子. 再对 V_{n-1} 仿此约化下去，不难得到

$$
\begin{aligned}
V_n &= \prod_{j=1}^{n-1}(x_n - x_j)V_{n-1} \\
&= \left[\prod_{j=1}^{n-1}(x_n - x_j)\right]\left[\prod_{j=1}^{n-2}(x_{n-1} - x_j)\right] \cdot V_{n-2} \\
&= \left[\prod_{j=1}^{n-1}(x_n - x_j)\right]\left[\prod_{j=1}^{n-2}(x_{n-1} - x_j)\right]\cdots\left[\prod_{j=1}^{2}(x_3 - x_j)\right] \\
&\quad \cdot (x_2 - x_1) \\
&= \prod_{1\leqslant j<i\leqslant n}(x_i - x_j).
\end{aligned}
$$

其实,这个结果还可按下列考虑更简单直接地得出.因为含 x_i 和含 x_j 的任何两列相同时,即当 $x_i=x_j$ 时,行列式为零;显然,行列式的值含有一切可能的因子 (x_i-x_j)(共 $C_2^n=\frac{1}{2}n(n-1)$项)之积,即 $\prod_{1\leqslant j<i\leqslant n}(x_i - x_j)$,由于每一因子均取前者 (x_i)时得到展开式中的一项为 $x_1^0x_2^1x_3^2\cdots x_n^{n-1}$,正好与主对角线之积相同,故该连乘积具有正确符号和系数.由此可见,

$$
V(x_1, x_2, \cdots, x_n) = \prod_{1\leqslant j<i\leqslant n}(x_i - x_j),
$$

它是反对称函数.

3.8.2 Jacobi 行列式

(1) 多元微积分学表明,通过 $w_i=w_i(v_1,\cdots,v_n)(i=1,\cdots,n)$作变量变换从 $w_1,\cdots,w_n$ 变至 $v_1,\cdots,v_n$ 时有

$$
dw_1\cdots dw_n = J dv_1\cdots dv_n,
$$

其中 J 是以$\frac{\partial w_i}{\partial v_k}$($i$ 为行标,k 为列标)为元素的行列式,称为 **Jacobi 行列式**,记作

$$
J = \frac{\partial(w_1,\cdots,w_n)}{\partial(v_1,\cdots,v_n)} = \begin{vmatrix} \frac{\partial w_1}{\partial v_1} & \cdots & \frac{\partial w_1}{\partial v_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial w_n}{\partial v_1} & \cdots & \frac{\partial w_n}{\partial v_n} \end{vmatrix} = \det\left(\frac{\partial w_i}{\partial v_k}\right).
$$

(2) 考虑通过复合函数

$$
w_i = w_i[v_j(u_k)] = w_i[v_1(u_1,\cdots,u_n),\cdots,v_n(u_1,\cdots,u_n)]
$$

作变量变换从 $w_1,\cdots,w_n\to v_1,\cdots,v_n\to u_1,\cdots,u_n$.按照复合函数求导法则有

$$
\frac{\partial w_i}{\partial u_k} = \sum_j \frac{\partial w_i}{\partial v_j}\frac{\partial v_j}{\partial u_k}.
$$

若令$\frac{\partial w_i}{\partial v_j}\equiv a_{ij}$,$\frac{\partial v_j}{\partial u_k}\equiv b_{jk}$,$\frac{\partial w_i}{\partial u_k}\equiv c_{ik}$,则有

$$c_{ik}=\sum_j a_{ij}b_{jk} \quad 或 \quad C=AB.$$

另一方面，多元微积分学给出

$$\begin{cases}\mathrm{d}w_1\cdots\mathrm{d}w_n=(\det C)\mathrm{d}u_1\cdots\mathrm{d}u_n,\\ \mathrm{d}w_1\cdots\mathrm{d}w_n=(\det A)\mathrm{d}v_1\cdots\mathrm{d}v_n,\\ \qquad\qquad\qquad \mathrm{d}v_1\cdots\mathrm{d}v_n=(\det B)\mathrm{d}u_1\cdots\mathrm{d}u_n;\end{cases}$$

由后两式立即得到

$$\mathrm{d}w_1\cdots\mathrm{d}w_n=(\det A)(\det B)\mathrm{d}u_1\cdots\mathrm{d}u_n,$$

所以，

$$\det C=\det(AB)=(\det A)(\det B),$$

这就证明了 3.4.2 节关于行列式之积的结果.

(3) 将上面的结果用 Jacobi 行列式的记号写出，可得下列重要公式：

$$\frac{\partial(w_1,\cdots,w_n)}{\partial(u_1,\cdots,u_n)}=\frac{\partial(w_1,\cdots,w_n)}{\partial(v_1,\cdots,v_n)}\frac{\partial(v_1,\cdots,v_n)}{\partial(u_1,\cdots,u_n)}.$$

(4) 由 $w_i=w_i(v_1,\cdots,v_n)(i=1,\cdots,n)$ 求微分可得

$$\mathrm{d}w_i=\sum_{k=1}^n\frac{\partial w_i}{\partial v_k}\mathrm{d}v_k,\quad i=1,\cdots,n;$$

由此可见，Jacobi 行列式 $\det\left(\frac{\partial w_i}{\partial v_k}\right)$ 即其系数矩阵的行列式；若不为零，则可解出 $(\mathrm{d}v_1,\cdots,\mathrm{d}v_n)$ 为 $(\mathrm{d}w_1,\cdots,\mathrm{d}w_n)$ 的函数. 因此有

$$\frac{\partial(w_1,\cdots,w_n)}{\partial(v_1,\cdots,v_n)}\frac{\partial(v_1,\cdots,v_n)}{\partial(w_1,\cdots,w_n)}=1 \quad 即$$

$$\left[\frac{\partial(w_1,\cdots,w_n)}{\partial(v_1,\cdots,v_n)}\right]^{-1}=\frac{\partial(v_1,\cdots,v_n)}{\partial(w_1,\cdots,w_n)}.$$

(5) 例如，从直角坐标 (x,y,z) 变换至柱面坐标 (ρ,φ,z) 和球面坐标 (r,θ,φ) 的 Jacobi 行列式分别为：

$$\frac{\partial(x,y,z)}{\partial(\rho,\varphi,z)}=\rho \quad 和 \quad \frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)}=r^2\sin\theta;$$

于是，体积元变换为

$$\mathrm{d}x\mathrm{d}y\mathrm{d}z=\rho\mathrm{d}\rho\mathrm{d}\varphi\mathrm{d}z=r^2\mathrm{d}r\sin\theta\mathrm{d}\theta\mathrm{d}\varphi.$$

3.8.3 Wroński 行列式

由一组 k 个函数 $y_1(x),\cdots,y_k(x)$ 及其至 $k-1$ 阶导数构成的下列行列式：

$$W(x)=\begin{vmatrix}y_1 & y_2 & \cdots & y_k\\ y_1' & y_2' & \cdots & y_k'\\ \vdots & \vdots & \ddots & \vdots\\ y_1^{(k-1)} & y_2^{(k-1)} & \cdots & y_k^{(k-1)}\end{vmatrix},$$

称为 **Wroński** 行列式，它在线性常微分方程中可用于判定 $y_1,\cdots,y_k$ 是否为其线性无关解.

3.9 行列式的导数与极限

3.9.1 行列式的导数

(1) 用行列式的导数表示代数余子式

由行列式的定义知 $|A|=|a_{rs}|$ 是 $a_{rs}(r=1,\cdots,n)$ 的一次齐式，于是，Euler 定理给出：

$$|A|=\sum_{r=1}^{n}a_{rs}\frac{\partial|A|}{\partial a_{rs}}=\left(\frac{\partial|A|}{\partial\boldsymbol{a}^{(s)}}\right)^{\mathrm{T}}\cdot\boldsymbol{a}^{(s)}=\sum_{r=1}^{n}a_{rs}A_{rs},$$

即 a_{rs} 的代数余子式 A_{rs} 可表示为

$$A_{rs}=\frac{\partial|A|}{\partial a_{rs}}.$$

因此，若在行列式 $|A|$ 中的 s 列 $\boldsymbol{a}^{(s)}$ 用 $\boldsymbol{c}$ 替换，则

$$\det A(\boldsymbol{a}^{(s)}\to\boldsymbol{c})=\det(\boldsymbol{a}'\cdots\boldsymbol{a}^{(s-1)}\ \boldsymbol{c}\,\boldsymbol{a}^{(s+1)}\cdots\boldsymbol{a}^{(n)})=\sum A_{rs}c_r.$$

(2) 函数的行列式

若矩阵 $A(\alpha)=\{a_{rs}(\alpha)\}$，则 $\frac{\mathrm{d}}{\mathrm{d}\alpha}A(\alpha)=\left\{\frac{\mathrm{d}}{\mathrm{d}\alpha}a_{rs}(\alpha)\right\}$.

对于行列式 $|A(\alpha)|$，则有

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}(\det A)=\sum_{s=1}^{n}\det A\left(\boldsymbol{a}^{(s)}\to\frac{\mathrm{d}}{\mathrm{d}\alpha}\boldsymbol{a}^{(s)}\right);$$

另一方面，还可求得

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\alpha}\ln|A|&=\frac{1}{|A|}\frac{\mathrm{d}|A|}{\mathrm{d}\alpha}=\frac{1}{|A|}\sum_r\sum_s\frac{\partial|A|}{\partial a_{rs}}\frac{\mathrm{d}a_{rs}}{\mathrm{d}\alpha}\\&=\frac{1}{|A|}\sum_r\sum_s A_{rs}\frac{\mathrm{d}a_{rs}}{\mathrm{d}\alpha}\\&=\sum_s\sum_r(A^{-1})_{sr}\left(\frac{\mathrm{d}}{\mathrm{d}\alpha}A\right)_{rs}\qquad[3.6.1\text{ 节}]\\&=\sum_s\left(A^{-1}\frac{\mathrm{d}}{\mathrm{d}\alpha}A\right)_{ss}\\&=\mathrm{tr}\left(A^{-1}\frac{\mathrm{d}}{\mathrm{d}\alpha}A\right);\end{aligned}$$

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}|A|=|A|\ \mathrm{tr}\left(A^{-1}\frac{\mathrm{d}}{\mathrm{d}\alpha}A\right).$$

3.9.2 行列式的极限

(1) 若 $\lim\limits_{\alpha\to\infty} b_{ik}(\alpha)=a_{ik}$，则 $\lim\limits_{\alpha\to\infty}|B(\alpha)|=|A|$.

证 因为这是有限数目(n 个)因子乘积的有限项数($n!$ 项)之代数和.

[参考：数学分析中证明有下列结果：

若 $\lim\limits_{\alpha\to\infty}x(\alpha)=X$，$\lim\limits_{\alpha\to\infty}y(\alpha)=Y$，则导致

$$\lim_{\alpha\to\infty}x(\alpha)y(\alpha)=XY \quad 及 \quad \lim_{\alpha\to\infty}[x(\alpha)+y(\alpha)]=X+Y.]$$

(2) 若 $|A|=0$，则有可能将 A 当作 $B(\alpha)$ 的极限，而对每个 α 有 $|B(\alpha)|\neq 0$.

证 考虑 $B(\alpha)$ 具有 $b_{ik}(\alpha)=a_{ik}+\lambda\delta_{ik}\left(\lambda=\dfrac{1}{\alpha}\right)$，于是

$$|B(\alpha)|=\lambda^n+\cdots+|A|;$$

它是 λ 的 n 次多项式，仅对 n 个 λ 值会变为零. 避开这些特定值，于是 $|B(\alpha)|\neq 0$，但是 $\lim\limits_{\alpha\to\infty}|B(\alpha)|=|A|=0$.

第 4 章　线性方程组的求解

第 1 章关于线性变换的研究和第 3 章关于行列式的知识，可以应用于线性方程组的求解.

4.1　引　　言

设含有 n 个未知量、n 个方程的方程组为

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = c_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = c_2, \\ \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = c_n, \end{array}\right\}$$

或者缩写为

$$\sum_{k=1}^{n} a_{ik}x_k = c_i \quad (i = 1,2,\cdots,n),$$

其中 a_{ik} 和 c_i 全为已知量，而 $x_1,x_2,\cdots,x_n$ 则是欲寻求的未知量. 用我们的语言来说，线性方程组的求解就是要寻找一个 $\boldsymbol{x}=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$，使得它的线性变换 $\boldsymbol{y}=A\boldsymbol{x}$ 能与 $\boldsymbol{c}$ 一致；这里 $\boldsymbol{y}=\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$，$\boldsymbol{c}=\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$，而 $A=\{a_{ik}\}$.

根据第 1 章 1.6.4 和 1.6.5 节的结果，我们有关于线性方程组求解的下述定理.

定理　若线性方程组的系数矩阵 $A=\{a_{ik}\}$ 是非奇异的，则不管常矢量 $\boldsymbol{c}$ 是怎么样的，方程组的解是唯一的；特别是，当且仅当 $\boldsymbol{c}=\boldsymbol{0}$ 时有平凡解 $\boldsymbol{x}=\boldsymbol{0}$. 若 A 是奇异的，当 $\boldsymbol{c}$ 位于子空间 $\boldsymbol{S}$[由列矢量集合 $\boldsymbol{a}',\cdots,\boldsymbol{a}^{(n)}$ ($\boldsymbol{a}^{(k)}=(a_1^{(k)}\cdots a_n^{(k)})^{\mathrm{T}}$，$a_i^{(k)}=a_{ik}$) 所生成]之外时无解，而当 $\boldsymbol{c}$ 位于 $\boldsymbol{S}$ 之内时有解但不唯一；$\boldsymbol{c}=\boldsymbol{0}$ 总是位于 $\mathbf{S}$ 之内，因为 $\boldsymbol{x}=\boldsymbol{0}\rightarrow\boldsymbol{y}=\boldsymbol{0}$.

因此，对于给定 A，只有两种可能：或者非齐次方程组($\boldsymbol{c}\neq\boldsymbol{0}$)具有唯一解，或者

齐次方程组($\boldsymbol{c}=\boldsymbol{0}$)具有若干非平凡解($\boldsymbol{x}\neq\boldsymbol{0}$).

记住这个一般观点,从而必须设计一个对一切情况都有效的实用求解方法.最实用的方法看来是为供数值工作使用而经标准化了的**消元法**[①];具体见下面的例子.

4.2 Gauss 消元法

用消元法求线性方程组的解,基本步骤是:将 n 个未知量逐一消去以求得一个未知量之解,然后代入仅有两个未知量的方程以求得另一未知量之解,以此类推直至求得所有未知量之解.

4.2.1 用消元法求数值解的例子

现举一例,用以说明消元法之具体运用.

$$4.17x-2.13y+1.17z+2.55=0,$$
$$-1.03x+3.71y+0.65z+1.15=0,$$
$$1.32x-1.06y+4.58z-2.11=0.$$

对于三位数计算,总是可以应用计算尺;而对于更加准确的计算,则需应用对数表或计算器或计算机.

对于数值求解,以明确方式具体安排计算工作是绝对必要的.具体安排要使计算工作相当方便,但更重要的是,应该容易进行验算以防备可能的误差或错误.

上述例子的数值求解的具体步骤,用列表进行计算并依次进行说明如下(见表 4.1):

(a) 注意事项

① 总是隔一行写.

② 增校验和 s 一列,只为校验;校验与计算同时进行.

③ 校验 s 列是否为前三列之和,否则有错误.

④ 校验 s 列是否为前二列之和,否则有错误.

(b) 填表次序

(i) 首先将原方程系数分别隔行填入 $a',a'',a''',-c$ 四列.

(ii) 将前四列之和填入 s 列.

(iii) 将(1)行乘以$\frac{1.03}{4.17}$或$-\frac{1.32}{4.17}$而得,分别填入(2′)及(3′)行.

(iv) 将所得行与原有行相加而得,分别填入Ⅱ段(4)及(5)行.

① C. Runge, H. König, *Vorlesungen über Numerisches Rechnen*, Springer, Berlin, 1924, Kap. 2.

表 4.1　具体计算步骤列表

		a'	a''	a'''	$-c$	$S(=a'+a''+a'''-c)$②			
Ⅰ	(1)	4.17(i)	−2.13(i)	1.17(i)	2.55(i)	5.76(ii)	3.28(ix)	4.87(xii)	0.00(xv)
	(1′)	①			0.73(viii)		1.59(xi)	−4.87(xiv)	$x=-1.168$(xiii)
	(2)	−1.03(i)	3.71(i)	0.65(i)	1.15(i)	4.48(ii)			
	(2′)	1.03(iii)	−0.53(iii)	0.29(iii)	0.63(iii)	1.42(iii)			
	(3)	1.32(i)	−1.06(i)	4.58(i)	−2.11(i)	2.73(ii)			
	(3′)	−1.32(iii)	0.67(iii)	−0.37(iii)	−0.81(iii)	−1.82(iii)			
Ⅱ	(4)		3.18(iv)	0.94(iv)	1.78(iv)	5.90(iv)③	2.37(ix)	0.00(xii)	
	(4′)				0.59(viii)		−2.37(xi)	$y=-0.745$(x)	
	(5)		−0.39(iv)	4.21(iv)	−2.92(iv)	0.91(iv)③			
	(5′)		0.39(v)	0.12(v)	0.22(v)	0.72(v)			
Ⅲ	(6)			4.33(vi)	−2.70(vi)	1.63(vi)④	0.00(ix)		
	(6′)				2.70(viii)		$z=0.624$(vii)		

(v) 将Ⅱ段(4)行乘以$\frac{0.39}{3.18}$而得,填入Ⅱ段(5′)行.

(vi) 将Ⅱ段(5)与(5′)行相加而得,填入Ⅲ段(6)行.

(vii) 由(6)行(第Ⅲ段第一行)求 z: $4.33z-2.70=0$, ∴ $z=0.624$;增一列.

(viii) $\boldsymbol{a}'''$列每段之第一行乘 $z=0.624$,所得填入$-\boldsymbol{c}$列.

(ix) $-\boldsymbol{c}$ 列每段之第一行两数相加,填入 $z=0.624$ 列.

(x) 由(4)行(第Ⅱ段第一行)求 y: $3.18y+2.37=0$, ∴ $y=-0.745$;增一列.

(xi) $\boldsymbol{a}''$列每段之第一行乘 $y=-0.745$,所得填入 $z=0.624$ 列.

(xii) 将 $z=0.624$ 列前两段两数相加,填入 $y=-0.745$ 列.

(xiii) 由第Ⅰ段第一行求 x: $4.17x+4.87=0$, ∴ $x=-1.168$, 增一列.

(xiv) $\boldsymbol{a}'$列第一行乘 $x=-1.168$,所得填入 $y=-0.745$ 列.

(xv) 将 $y=-0.745$ 列的前两数相加,所得填入 $x=-1.168$ 列.

(xvi)(xvii) 最后将求得之解代入原方程并进行验算(见表 4.2).

表 4.2 最后验算表

	方程(1)	方程(2)	方程(3)	
$\boldsymbol{a}'x$:	−4.87(xvi)	1.20(xvi)	−1.54(xvi)	
$\boldsymbol{a}''y$:	1.59(xvi)	−2.76(xvi)	0.79(xvi)	
$\boldsymbol{a}'''z$:	0.73(xvi)	0.41(xvi)	2.86(xvi)	+
与 $\boldsymbol{c}$ 验算	−2.55(xvii)	−1.15(xvii)	2.11(xvii)	

备注 $\boldsymbol{a}'$列之第一行,$\boldsymbol{a}''$列第二段第一行,等等,数值必须为**最大**者(不计+、−号),称为**主元**;此法即称为**全主元消元法**. 否则的话,会牺牲准确度;因此,格式可作相当变化. 这里所谓最大,包括准确度而言. 比如,4.17>1.03,但 4.17<1.030,盖因 1.030 方程可统以 10 倍之而得 10.30 之故也.

4.2.2 关于数值解的讨论

(1) 系数方阵是奇异阵时的情形

在消元过程的某一步若出现 $0x+0y+\cdots=c$ 时,则表明系数方阵是奇异的这种情况;在此情况下,若 $c\neq0$ 则表明方程组无解,而 $c=0$ 时则表明有多组解.

(2) 非齐次项是代数量时给出逆变换

若$-\boldsymbol{c}$ 列是代数量时,同样的计算给出逆变换,如果它存在的话;并且也对非奇异或奇异情况提供一种检验.

(3) 数值解的准确度问题

在数值求解的情况,$0x+0y+\cdots=c$ 中出现的往往是"近似零". 比如 $0.00y-2.38=0$ 导致 $y<-238$, $238\leqslant y$,因为"近似的"0.00 在±0.01 之间变化. 因此,如

果 a_{ik} 是以近似数值给定，我们并不能确切地知道该变换是否是奇异的.

(4) 系数方阵为奇异阵的进一步讨论

通常，由于要求 $A\boldsymbol{x}=\boldsymbol{0}$ 应具有非平凡解 $\boldsymbol{x}\neq\boldsymbol{0}$，则肯定会出现奇异情况. 这要求 A 为奇异方阵，并且足以在一类线性变换中来挑选 A，而不是先前那样限于确指的一个.

例如，设 A 为非奇异方阵，我们要求 $A\boldsymbol{x}$ 为 $\boldsymbol{x}$ 的标量倍数（$A\boldsymbol{x}$ 平行于 $\boldsymbol{x}$），比如 $\lambda\boldsymbol{x}(\lambda\neq0)$，$\lambda$ 尚待确定. 于是，

$$A\boldsymbol{x}=\lambda\boldsymbol{x} \quad \text{或} \quad [A-\lambda I]\boldsymbol{x}=\boldsymbol{0},$$

其中 I 是单位方阵；或者写成

$$A(\lambda)\boldsymbol{x}=\boldsymbol{0}, \quad \text{而} \quad A(\lambda)\equiv A-\lambda I.$$

当 λ 具有这样一些值，它们使得 $A(\lambda)$ 是奇异阵时，这种情况会出现.

这称为本征值问题，将在第 6 章 6.3 节进行讨论.

4.3 Cramer 法则

(1) 系数矩阵是非奇异的

对于线性方程组

$$\sum_k a_{ik}x_k=c_i \quad \text{或} \quad A\boldsymbol{x}=\boldsymbol{c},$$

若 $\det A\neq0$，则用 a_{is} 的代数余子式 A_{is} 乘前式并对 i 求和，以及注意到（见第 3 章 3.3.2 节）

$$\sum_i a_{ik}A_{is}=\delta_{ks}(\det A), \quad \sum_i c_iA_{is}=\det A(\boldsymbol{a}^{(s)}\to\boldsymbol{c}),$$

立即得到唯一解

$$x_s=\frac{\sum_i c_iA_{is}}{\det A}=\frac{\det A(\boldsymbol{a}^{(s)}\to\boldsymbol{c})}{\det A};$$

这也可由 $\boldsymbol{x}=A^{-1}\boldsymbol{c}$ 直接得出. 这就是线性方程组求解的 **Cramer 法则**.

可以看出，若需在若干不同列矢量 $\boldsymbol{c}$ 的情况下求解 $\boldsymbol{x}$，一次求出 A^{-1} 后，只要执行乘法就可求得各个解.

但是应该指出，这只有在 n 不大的情况下才可行.

(2) 系数矩阵与增广矩阵同秩

对于系数矩阵 $A=\{\boldsymbol{a}^{(1)},\cdots,\boldsymbol{a}^{(n)}\}$ 是奇异阵或长方阵 $m\neq n$ 的情况，若矩阵 A 与其增广矩阵

$$\overline{A}=\{\boldsymbol{a}^{(1)},\cdots,\boldsymbol{a}^{(n)},\boldsymbol{c}\}$$

为同秩，即

$$\text{rank}A = \text{rank}\overline{A} = r \leqslant \min(n,m),$$

方程组有解，但不唯一.

不失一般性，设

$$\det A_r = \begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \neq 0,$$

则原方程组与仅含前 r 个方程的方程组同解. 将后者改写为

$$\sum_{k=1}^{r} a_{ik} x_k = c_i' = c_i - \sum_{j=r+1}^{n} a_{ij} x_j \quad (i = 1,\cdots,r),$$

于是可应用前述 Cramer 法则求解，得到

$$x_s = \frac{\det A_r(\boldsymbol{a}^{(s)} \to \boldsymbol{c}')}{\det A_r} \quad (s = 1,\cdots,r).$$

显然，$x_1,\cdots,x_r$ 是 $(n-r)$ 个自由未知量 $x_{r+1},\cdots,x_n$ 的线性函数，因为后者可取任意值，所以解不唯一.

另一方面，若 A 与 $\overline{A}$ 不同秩，则为矛盾方程组而无解.

4.4 迭 代 法①②③

前两节的方法，在无舍入误差情况下，经过有限次算术运算即可求得精确解；它们称为**直接法**. 许多实际问题最后常归结为一个或一些大型**稀疏矩阵**的线性方程组的求解；这里方程组的未知量成千上万，但其系数矩阵的非零元素所占比例很小或分布很有规律. 对于此类方程组，直接法无能为力，一般采用**迭代法**求解. 迭代法采取逐次逼近的方法，从一个起步解 $\boldsymbol{x}^{(0)}$ 出发，按照一定迭代公式构造一个解的收敛序列 $\boldsymbol{x}^{(0)},\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\cdots$，直至求得满足一定精度要求的近似解为止.

4.4.1 几种常用迭代法

下面介绍三种常用的迭代法.

(1) Jacobi 迭代法

对于线性方程组

$$\sum_{k=1}^{n} a_{ik} x_k = c_i \quad (i = 1,\cdots,n),$$

① 参考：《中国大百科全书・数学》，624 页，松弛法（胡家赣撰），中国大百科全书出版社，北京，1988.

② 参考：胡家赣，《线性代数方程组的迭代解法》，科学出版社，北京，1988.

③ 参考：D. M. Young, *Iterative Solution of Large Linear System*, Academic Press, New York, 1971.

若 $a_{ii}(i=1,\cdots,n)$ 均不为零，则由迭代公式

$$x_i^{(s+1)} = a_{ii}^{-1}\Big[-\sum_{k\neq i} a_{ik}x_k^{(s)} + c_i\Big] \quad (i = 1,\cdots,n)$$

和给定的初值 $x_1^{(0)},\cdots,x_n^{(0)}$，就可求出 $x_1^{(1)},\cdots,x_n^{(1)}$，如此继续直到满足计算精度要求为止，这种计算格式称为 **Jacobi 迭代法**.

(2) Gauss-Seidel 迭代法

若在 Jacobi 迭代法中的每一步均将已迭代出的 $x_k^{(s+1)}(s=0,1,2,\cdots;k=1,2,\cdots,i-1)$ 立即代入其后各方程中，则得到迭代公式

$$x_i^{(s+1)} = a_{ii}^{-1}\Big[-\sum_{k<i} a_{ik}x_k^{(s+1)} - \sum_{k>i} a_{ik}x_k^{(s)} + c_i\Big]$$
$$(i = 1,2,\cdots,n),$$

它称为 **Gauss-Seidel 迭代法**.

(3) 逐次超松弛迭代法

假定对 $(s+1)$ 次近似 $\boldsymbol{x}^{(s+1)}$ 已知其中分量 $x_k^{(s+1)}(k=1,2,\cdots,i-1)$，可按上述 Gauss-Seidel 迭代公式定义一辅助量 $\tilde{x}_i^{(s+1)}$，而 $x_i^{(s+1)}$ 则通过 $x_i^{(s)}$ 与 $\tilde{x}_i^{(s+1)}$ 的某种求平均予以确定，即

$$\begin{aligned} x_i^{(s+1)} &= (1-\omega)x_i^{(s)} + \omega\,\tilde{x}_i^{(s+1)} \\ &= x_i^{(s)} + \omega(\tilde{x}_i^{(s+1)} - x_i^{(s)}) \quad (\omega \in \mathbf{R}), \end{aligned}$$

将 $\tilde{x}_i^{(s+1)}$ 的具体形式代入，可得下列迭代公式

$$x_i^{(s+1)} = x_i^{(s)} + \omega a_{ii}^{-1}\Big[-\sum_{k<i} a_{ik}x_k^{(s+1)} - a_{ii}x_i^{(s)} - \sum_{k>i} a_{ik}x_k^{(s)} + c_i\Big]$$
$$(i = 1,2,\cdots,n;s = 0,1,2,\cdots);$$

它称为**逐次超松弛迭代法**(英文缩写为 SOR). 这里 ω 称为**松弛因子**，若 $\omega>1$ 为**超松弛**而 $\omega<1$ 为**低松弛**. 显而易见，$\omega=1$ 时即为 Gauss-Seidel 迭代法.

4.4.2　迭代格式的矩阵形式

将线性代数方程组

$$A\boldsymbol{x} = \boldsymbol{c}$$

的系数矩阵 A 分裂成

$$A = D - L - U,$$

其中

$$\begin{aligned} D &= \mathrm{diag}[a_{11},\cdots,a_{nn}], \\ L &= \{l_{ij} = -a_{ij}(i>j), l_{ij} = 0(i\leqslant j)\}, \\ U &= \{u_{ij} = -a_{ij}(i<j), u_{ij} = 0(i\geqslant j)\}, \end{aligned}$$

分别为对角矩阵，严格下三角形矩阵和严格上三角形矩阵. 于是，三种迭代格式的矩阵形式可表达为

$$\begin{aligned}
&\text{J} && Dx^{(s+1)} = (L+U)x^{(s)} + c,\\
&\text{GS} && (D-L)x^{(s+1)} = Ux^{(s)} + c,\\
&\text{SOR} && (D-\omega L)x^{(s+1)} = [(1-\omega)D+\omega U]x^{(s)} + \omega c.
\end{aligned}$$

它们可以写成下列统一形式：

$$x^{(s+1)} = Gx^{(s)} + g \quad (s = 0,1,2,\cdots),$$

其中 G 称为**迭代矩阵**. 对于三种迭代法 G 和 g 的具体形式分别为

J　$G=J=D^{-1}(L+U)$，　　$g=D^{-1}c$；

GS　$G=H=(D-L)^{-1}U$，　　$g=(D-L)^{-1}c$；

SOR　$G=H(\omega)=(D-\omega L)^{-1}[(1-\omega)D+\omega U]$，　　$g=(D-\omega L)^{-1}\omega c$.

4.4.3 迭代收敛性

对于迭代法，必须考虑其**收敛性**的问题，这里不加证明地引述下列定理.

迭代格式

$$x^{(s+1)} = Gx^{(s)} + g$$

收敛的充要条件为

$$\rho(G) < 1.$$

这里 $\rho(G)$ 表示 n 阶方阵 G 的**谱半径**，定义为

$$\rho(G) = \max_{1\leqslant i\leqslant n} |\lambda_i|,$$

其中 $\lambda_i(i=1,\cdots,n)$ 为 G 的本征值(关于本征值，见第 6 章 6.3 节).

对于任意一种**矩阵范数** $\|A\|$，有(见第 3 章 3.6.4 节之注)

$$\rho(A) \leqslant \|A\|.$$

所以，上述迭代格式收敛的一个充分条件为

$$\|G\| < 1.$$

还可定义渐近收敛速率

$$R(G) = \lim_{s\to\infty}\left[-\frac{1}{s}\lg\|G^s\|\right] = -\lg\rho(G).$$

若需提高精度 10^{-m}，则约需迭代次数为

$$s = m/R(G).$$

对于**对角优势矩阵**，即具有

$$|a_{ii}| > \sum_{k\neq i}|a_{ik}| \quad (i=1,\cdots,n) \quad \text{或}$$

$$|a_{kk}| > \sum_{i\neq k}|a_{ik}| \quad (k=1,\cdots,n)$$

的矩阵 $A=\{a_{ik}\}$，Jacobi 迭代法和 Gauss-Seidel 迭代法收敛，而逐次超松弛迭代法在

$$0 < \omega < 2/[1+\rho(|J|)]$$

范围内收敛. 这里 $J=D^{-1}(L+U)$，而 $|J|$ 表示矩阵 J 的元素取绝对值所成的矩阵.

关于以上有关定理的证明以及其他相关情况，可参考前引专著.

4.4.4 松弛因子的选取

如何选取最优的 ω 使松弛法收敛最快是一个困难问题.

若矩阵 $A=D-L-U$ 具有性质使矩阵

$$J(\alpha) = D^{-1}(\alpha L + \alpha^{-1}U) \quad (\alpha \in \mathbf{C}, \alpha \neq 0)$$

的本征值 μ 与 α 无关，则矩阵 A 被称为**相容次序矩阵**.

对于相容次序矩阵，可以证明，J 的本征值 μ 与 SOR 法的迭代矩阵 $H(\omega)$ 的本征值 $\lambda(\omega)$ 之间有下列关系：

$$[\lambda(\omega)+\omega-1]^2 = \lambda(\omega)\omega^2\mu^2.$$

特别是当 $\omega=1$ 时有 $\lambda=\mu^2$，即有

$$\rho(H) = \rho(J)^2,$$

其中 H 是 GS 法的迭代矩阵. 由于迭代收敛速率与 ρ 直接相关，要达到相同精度，Jacobi 法所需迭代步约为 GS 法的二倍.

最优松弛因子 ω_b 由

$$\rho(H(\omega_b)) = \min_{\omega \in \mathbf{R}} \rho(H(\omega)) = \min_{0<\omega<2} \rho(H(\omega))$$

表征. 对于相容次序矩阵，可以求得

$$\omega_b = \frac{2}{1+\sqrt{1-\rho(J)^2}},$$

$$\rho(\omega_b) = \omega_b - 1 = \left(\frac{\rho(J)}{1+\sqrt{1-\rho(J)^2}}\right)^2.$$

前引 D. M. Young 的专著 177—185 页表列了 ω_b 对 $\rho(J)$ 的函数值.

4.4.5 一个例子

用五点差分格式解二维 Laplace 方程

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

若网格间距 $h=\frac{1}{3}$ 和边界条件取为

$$u(x,1) = 4500x(1-x), \quad u(1,y) = u(x,0) = u(0,y) = 0$$

时，得到下列线性方程组

$$\begin{bmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1000 \\ 1000 \end{bmatrix}.$$

系数矩阵是对角优势矩阵，也是相容次序矩阵. 容易求得 J 的本征值是 $0,0,\frac{1}{2}$，$-\frac{1}{2}$，所以有（一般有 $\rho(J)=\cos \pi h$）

$$\rho(J)=\frac{1}{2}<1,\quad \rho(H)=\frac{1}{4}<1,$$

$$\omega_b=\frac{2}{1+\sqrt{3}/2}=1.0718,\quad \rho(\omega_b)=0.0718<1,$$

因此，三种迭代法均收敛.

对于起步解 $\boldsymbol{x}^{(0)}=\boldsymbol{0}$，Jacobi 法经 19 步迭代收敛至 6 位有效数字，Gauss-Seidel 法只需 12 步，而取 $\omega_b=1.0718$ 时 SOR 法仅需 7 步. 注意到，若应用网格间距 $h=\frac{1}{80}$，未知量增加至 $\left(\frac{1}{h}-1\right)^2=79^2$，则 Jacobi 法收敛至同样精度需经 18 000 步迭代，GS 法约减为一半，而 SOR 法收敛速率比 GS 法约快 $2/(\pi h)\approx 51$ 倍.

习　题

4.1　用(a) Gauss 消元法和(b) Cramer 法则分别求下列方程组的数值解：

$$\begin{aligned} 2.86x-4.78y+3.19z&=1.375,\\ 6.32x+2.47y-5.64z&=4.12,\\ 3.03x-5.91y+1.553z&=1.362. \end{aligned}$$

(b)中行列式的计算采用凝聚法（见 3.3.4 节）.

4.2　求下列方程组的数值解[①②]：

$$\begin{bmatrix} 0.000\,003 & 0.213\,472 & 0.332\,147 \\ 0.215\,512 & 0.375\,623 & 0.476\,625 \\ 0.173\,257 & 0.663\,257 & 0.625\,675 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0.235\,262 \\ 0.127\,653 \\ 0.285\,321 \end{bmatrix}.$$

解　用 6 位数计算器，采用**全主元消元法**，选 0.215 512 为主元，消元法第一步得

① 参考：J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon, Oxford, 1965, p. 216.

② 参考：转引自 A. S. Deif, *Advanced Matrix Theory for Scientists and Engineers*, p. 61.

$$\begin{bmatrix} 0.215\,512 & 0.375\,623 & 0.476\,625 \\ 0 & 0.213\,467 & 0.332\,140 \\ 0 & 0.361\,282 & 0.242\,501 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0.127\,653 \\ 0.235\,260 \\ 0.182\,697 \end{bmatrix}.$$

再次选 0.361 282 为主元，消元法第二步得

$$\begin{bmatrix} 0.215\,512 & 0.375\,623 & 0.476\,625 \\ 0 & 0.361\,282 & 0.242\,501 \\ 0 & 0 & 0.188\,856 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0.127\,653 \\ 0.182\,697 \\ 0.127\,312 \end{bmatrix}.$$

回代求得

$$z = 0.674\,122, \quad y = 0.053\,2050, \quad x = -0.991\,291.$$

若将方程中数值看作准确值，至 10 位数的结果是

$$z = 0.674\,121\,4694, \quad y = 0.053\,203\,933\,91,$$
$$x = -0.991\,289\,4252.$$

由此可见，上述计算相当准确.

然而，如果不对主元进行选择，分别得

$$\begin{bmatrix} 0.000\,003 & 0.213\,472 & 0.332\,147 \\ 0 & -15\,334.9 & -23\,860.0 \\ 0 & 12\,327.8 & -19\,181.7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0.235\,262 \\ -16\,900.5 \\ -13\,586.6 \end{bmatrix},$$

$$\begin{bmatrix} 0.000\,003 & 0.213\,472 & 0.332\,147 \\ 0 & -15\,334.9 & -23\,860.0 \\ 0 & 0 & -0.500\,000 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0.235\,262 \\ -16\,900.5 \\ -0.200\,00 \end{bmatrix},$$

$$z = 0.400\,000, \quad y = 0.479\,723, \quad x = -1.333\,33;$$

结果非常差.

由上述对比可以看出，消元法中主元选择的极端重要性.

第 5 章　矢量与张量分析

5.1　矢量与张量的定义

我们知道，一个矢量 $\boldsymbol{x}$ 可通过非奇异线性变换 A 而变为另一个矢量 $\boldsymbol{y}$，用数学式表示，即

$$\boldsymbol{x} \to \boldsymbol{y} = A\boldsymbol{x};$$

这种变换同时具有加性和齐性，即，若

$$\boldsymbol{x}' \to \boldsymbol{y}', \quad \boldsymbol{x}'' \to \boldsymbol{y}'',$$

则有

$$\boldsymbol{x}' + \boldsymbol{x}'' \to \boldsymbol{y}' + \boldsymbol{y}'', \qquad \text{（加性）}$$

$$\lambda\boldsymbol{x}' \to \lambda\boldsymbol{y}', \qquad \text{（齐性）}$$

因而称为线性的.

数学家及物理学家总喜欢对已有概念加以推广，他们将矢量推广为这样一种数学量，它们当坐标变换（所谓坐标变换，即当 $\boldsymbol{x}\to\boldsymbol{y}=A\boldsymbol{x}$ 时，可将非奇异线性变换 A 想象为代表坐标变换）时，从一个量变成另一个量，这一变化仍然具有加性和齐性，即仍具有线性性质. 数学家将这种数学量称为**张量**，而**矢量**即为张量的简单情形. 主要观念是：矢量与张量均由其相对于某类非奇异齐次线性变换群的变换性质予以定义.

用数学式表示

$$\overset{(x)}{\Phi} \to \overset{(y)}{\Phi} = \overset{(A)}{\mathscr{A}}\overset{(x)}{\Phi},$$

这里变换系数 $\mathscr{A}$ 由坐标变换 A 确定，即其分量为 a_{ik} 的有限次积. 这种变换是非奇异线性变换，即当变换时，n 维空间还是变到 n 维空间（换句话说，只是坐标变换）. 而且还具有下列性质：若

$$\overset{(x)}{\Phi}{}' \to \overset{(y)}{\Phi}{}', \quad \overset{(x)}{\Phi}{}'' \to \overset{(y)}{\Phi}{}'',$$

则有

$$\overset{(x)}{\Phi}{}' + \overset{(x)}{\Phi}{}'' \to \overset{(y)}{\Phi}{}' + \overset{(y)}{\Phi}{}'', \qquad \text{（加性）}$$

$$\lambda\overset{(x)}{\Phi} \to \lambda\overset{(y)}{\Phi}, \qquad \text{（齐性）}$$

即变换具有线性性质.

5.2 Descartes 张量

Descartes 张量，即直角坐标系中的张量，它是 Euclid 空间中相对于正交群 $\mathbf{O}_n(\mathbf{R})$ 而定义的张量. 几何意义上说，当坐标系作正交变换时，张量具有加性和齐性，即具有线性性质. 换句话说，张量方程的形式不因坐标变换而改变，也就是与坐标的选择无关.

5.2.1 正交变换

对于正交变换(参见 2.5 节)

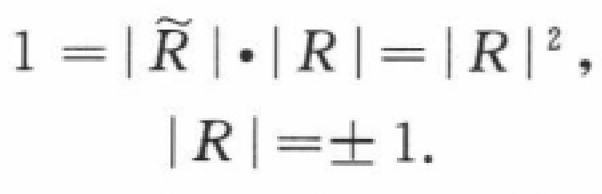
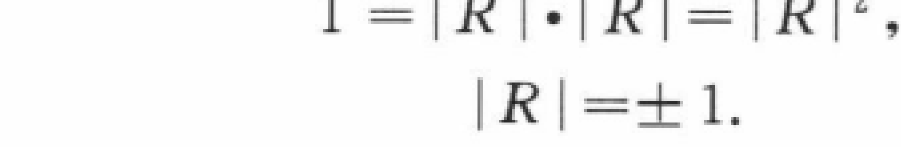
$$\boldsymbol{x} \to \boldsymbol{y} = R\boldsymbol{x},$$

$$\widetilde{R}R = R\widetilde{R} = 1 \quad 或 \quad R^{-1} = \widetilde{R};$$

可以证明

$$\det R = |R| = \pm 1.$$

因为对 $1 = \widetilde{R}R$ 取其行列式，得到

$$1 = |\widetilde{R}| \cdot |R| = |R|^2,$$

$$|R| = \pm 1.$$

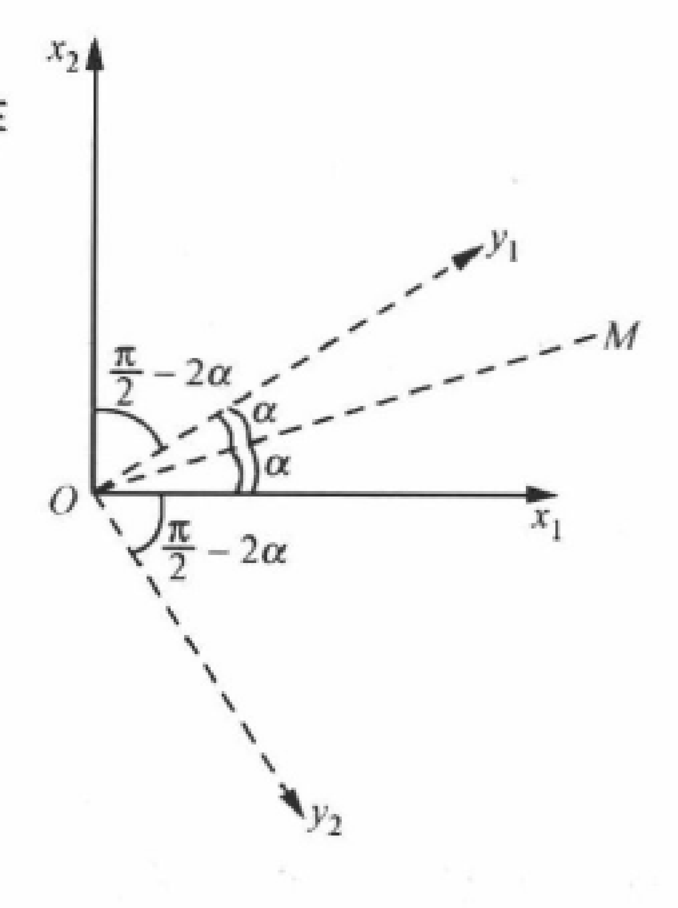
其中 $|R| = -1$ 代表反射，$|R| = +1$ 代表转动.

例如，坐标系绕 x_3 轴(垂直于纸面向上)，**转动** θ 角(如左图)的变换矩阵为

$$R_{\text{rot}} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad |R_{\text{rot}}| = 1; \quad \begin{cases} y_1 = x_1\cos\theta + x_2\sin\theta, \\ y_2 = -x_1\sin\theta + x_2\cos\theta, \\ y_3 = x_3. \end{cases}$$

注意，使矢量转动 θ 角的变换矩阵为 $R_{\text{rot}}^{\mathrm{T}}$.

另外，坐标系对 MOx_3 平面(与 x_1x_3 平面夹角 α)作**反射**(如右图)的变换矩阵为



$$R_{\text{ref}} = \begin{bmatrix} \cos2\alpha & \sin2\alpha & 0 \\ \sin2\alpha & -\cos2\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$|R_{\text{ref}}| = -1;$$

$$\begin{cases} y_1 = x_1\cos2\alpha + x_2\sin2\alpha, \\ y_2 = x_1\sin2\alpha - x_2\cos2\alpha, \\ y_3 = x_3. \end{cases}$$

同时注意到，$R_{\mathrm{ref}}(\alpha)R_{\mathrm{ref}}(\beta)=R_{\mathrm{rot}}(2(\beta-\alpha))$，即两次反射等价于一次转动.

5.2.2 Descartes 张量

(1) 一阶 Descartes 张量——矢量

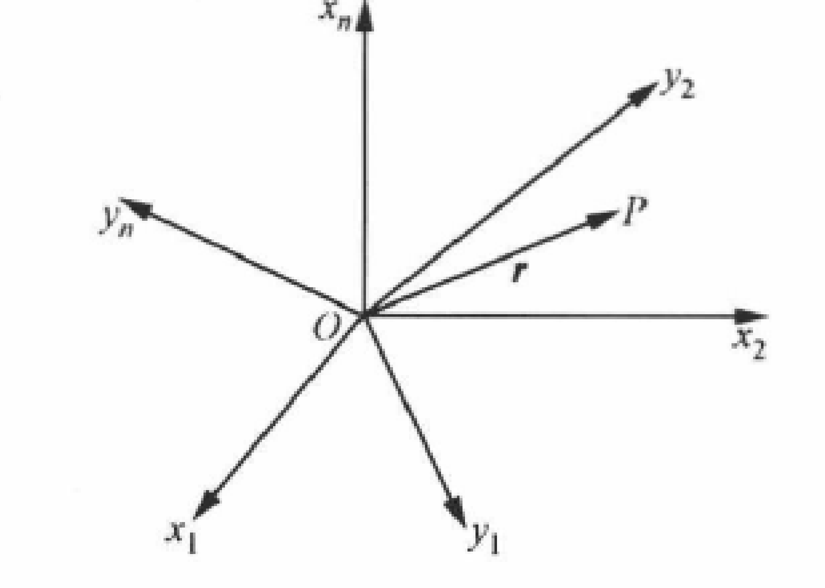
对于坐标变换(为了方便，下面直至 5.4 节均用 A 表示正交变换)

$$\boldsymbol{x}\to\boldsymbol{y}=A\boldsymbol{x}\quad 或\quad y_i=\sum_k a_{ik}x_k,$$

Descartes 一阶张量 $\boldsymbol{V}$ 的变换规则与之相像

$$\overset{(x)}{V}_k\to\overset{(y)}{V}_i=\sum_k a_{ik}\overset{(x)}{V}_k,$$

故亦称**矢量**.



位矢 $\boldsymbol{r}$(参见右图)

$$\overset{(x)}{\boldsymbol{r}}=(x_1,\cdots,x_n)^{\mathrm{T}},$$

$$\overset{(y)}{\boldsymbol{r}}=(y_1,\cdots,y_n)^{\mathrm{T}},$$

满足 $y_i=\sum_k a_{ik}x_k$，它是 Descartes 一阶张量的原型.

(2) 二阶 Descartes 张量

二阶张量的变换规则是

$$\overset{(y)}{\Phi}_{ij}=\sum_{kl}a_{ik}a_{jl}\overset{(x)}{\Phi}_{kl},$$

它共有 n^2 个分量.

并矢 $\boldsymbol{r}'\boldsymbol{r}''$是二阶 Descartes 张量的原型，因为

$$\overset{(x)}{\Phi}_{kl}=(\overset{(x)}{\boldsymbol{r}'}\overset{(x)}{\boldsymbol{r}''})_{kl}=(\overset{(x)}{\boldsymbol{r}'})_k(\overset{(x)}{\boldsymbol{r}''})_l=x'_kx''_l,$$

$$\overset{(y)}{\Phi}_{ij}=y'_iy''_j=\Big(\sum_k a_{ik}x'_k\Big)\Big(\sum_l a_{jl}x''_l\Big)$$

$$=\sum_{kl}a_{ik}a_{jl}x'_kx''_l=\sum_{kl}a_{ik}a_{jl}\overset{(x)}{\Phi}_{kl}.$$

(3) 三阶及高阶 Descartes 张量

三阶张量的变换规则是

$$\overset{(y)}{\Phi}_{ijk}=\sum_{pqr}a_{ip}a_{jq}a_{kr}\overset{(x)}{\Phi}_{pqr},$$

它共有 n^3 个分量. 更高阶张量可依此类推.

(4) 零阶张量——标量

零阶张量称为**标量**，有

$$\overset{(y)}{\varphi}=\overset{(x)}{\varphi},$$

即在坐标变换下保持其数值不变，故又称**不变量**.

例如，位矢 $\boldsymbol{r}$ 长度的平方是其原型，因为在正交变换下有

$$\sum_i y_i^2 = \sum_i x_i^2.$$

综上所述可以看出，标量、矢量和张量仅在量上不同，质上并无不同.

5.2.3 Descartes 张量的例子

(1) 应力与应变

应力定义为作用于固体单位面积上的力，它涉及面元法线与力两矢量之方向，需用二阶张量表示. 如在 Descartes 坐标系中，τ_{ij} 表示作用于与 x_i 轴垂直的面元上沿 j 方向的应力.

物体中的应力产生**应变或形变**. 形变前后弧元 $\mathrm{d}s_0$ 和 $\mathrm{d}s$ 之间有下列关系

$$\begin{aligned}\mathrm{d}s^2 = \sum_k \mathrm{d}x_k^2 &= \sum_{i,j,k} \frac{\partial x_k}{\partial x_{oi}} \frac{\partial x_k}{\partial x_{oj}} \mathrm{d}x_{oi} \mathrm{d}x_{oj} \\ &= \mathrm{d}s_0^2 + 2\sum_{i,j} e_{ij} \mathrm{d}x_{oi} \mathrm{d}x_{oj},\end{aligned}$$

其中所引进的二阶对称张量 e_{ij} 称为**应变张量**.

在小形变时，应力与应变之间有 **Hooke 定律**

$$\tau_{ij} = \sum_{k,l} c_{ijkl} e_{kl},$$

这里引进的四阶张量 c_{ijkl} 称为**弹性张量**. 可以证明，由于对称性，最多可有 21 个独立分量；而对具有立方对称的晶体，则仅有 3 个独立分量.

(2) 介电极化率和压电张量

当有电场 $\boldsymbol{E}$ 作用于晶体时，在其中产生电极化. 对于各向异性固体，在弱场下，电极化强度 $\boldsymbol{P}$ 与电场强度 $\boldsymbol{E}$ 之间有下列关系

$$P_i = \sum_j \chi_{ij} E_j,$$

其中 χ_{ij} 是二阶张量，称为**介电极化率张量**.

压电晶体在外应力作用下会产生电极化. 在弱应力下，电极化强度 P_i 与应力张量 τ_{jk} 之间有下列关系

$$P_i = \sum_{j,k} d_{ijk} \tau_{jk},$$

其中 d_{ijk} 是三阶张量，称为**压电张量**；一般仅有 18 个分量.

(3) 转动惯量

质点系作为刚体转动的总角动量为

$$\begin{aligned}\boldsymbol{L} &= \sum_a m_a [\boldsymbol{r}_a \times (\boldsymbol{\omega} \times \boldsymbol{r}_a)] \\ &= \sum_a m_a [r_a^2 \boldsymbol{\omega} - \boldsymbol{r}_a (\boldsymbol{r}_a \cdot \boldsymbol{\omega})],\end{aligned}$$

其中 m_a 是 a 质点的质量，$\boldsymbol{r}_a$ 是其相对于质心的位矢，而 $\boldsymbol{\omega}$ 是角速度；或者写成分量形式为

$$L_i = \sum_j I_{ij}\omega_j,$$

其中 I_{ij} 是**转动惯量张量**

$$I_{ij} = \sum_a m_a[\boldsymbol{r}_a^2\delta_{ij} - (\boldsymbol{r}_a)_i(\boldsymbol{r}_a)_j],$$

这里 δ_{ij} 是 Kronecker δ 符号.

5.3 Descartes 张量的运算

5.3.1 张量的线性相加

两个同阶张量 Φ 和 Ψ 可以线性相加

$$\lambda\Phi + \mu\Psi = \Theta,$$

结果所得 Θ 为一同阶新张量，这里 λ 和 μ 是标量系数.

证 张量的线性相加指相应分量的线性相加，因为张量当坐标变换时具有线性性质，所以 $\lambda\Phi+\mu\Psi$ 也是一个张量. 例如

$$\lambda \overset{(y)}{\Phi}_{ij\cdots k} + \mu \overset{(y)}{\Psi}_{ij\cdots k} = \sum_{p,q,\cdots,r} a_{ip}a_{jq}\cdots a_{kr}(\lambda \overset{(x)}{\Phi}_{pq\cdots r} + \mu \overset{(x)}{\Psi}_{pq\cdots r}),$$

$$\text{即}\ \overset{(y)}{\Theta}_{ij\cdots k} = \sum_{p,q,\cdots,r} a_{ip}a_{jq}\cdots a_{kr} \overset{(x)}{\Theta}_{pq\cdots r}.$$

5.3.2 张量的相等

两个同阶张量，当且仅当其所有对应分量均相等时，则称此**两张量相等**；即

$$\Phi_{ij\cdots k} = \Psi_{ij\cdots k}.$$

5.3.3 零张量

各阶张量都有全部分量均为零的**零张量**. 例如，

一阶零张量 $\mathbf{0}=(0,\cdots,0)^{\mathrm{T}}$ 或 $\overset{(x)}{0}_p=0$，

二阶零张量 $\overset{(x)}{0}_{pq}=0$，

r 阶零张量 $\overset{(x)}{0}_{p_1\cdots p_r}=0$.

同时注意到，

$$\overset{(y)}{0}_{i_1\cdots i_r} = \sum_{p_1\cdots p_r} a_{i_1p_1}\cdots a_{i_rp_r} \overset{(x)}{0}_{p_1\cdots p_r} = 0,$$

即，若一张量在任一坐标系中的全部分量为零，则它在一切坐标系中的全部分量均为零.

5.3.4　单位张量

二阶单位张量定义为

$$\overset{(x)}{I}_{kl} = \delta_{kl},$$

在三维空间中可写为

$$\mathsf{I} = \boldsymbol{ii} + \boldsymbol{jj} + \boldsymbol{kk},$$

其中 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 为在 x、y、z 轴方向上的**单位矢**；故又称**单位并矢**. 显然有

$$\begin{aligned}\overset{(y)}{I}_{ij} &= \sum_{k,l} a_{ik} a_{jl} \overset{(x)}{I}_{kl} = \sum_{k,l} a_{ik} a_{jl} \delta_{kl} \\ &= \sum_{k} a_{ik} a_{jk} = (AA^{\mathrm{T}})_{ij} = \delta_{ij},\end{aligned}$$

即在 (x) 坐标系中的 δ_{kl} 变换为 (y) 坐标系中的 δ_{ij}.

例如，压强就可用单位张量来表达，因为法线方向为 x 轴的平面只有 x 轴方向才有压强 p，其他方向都等于零. 所以

$$\mathsf{p} = p\mathsf{I},$$

p 为压力张量.

5.3.5　张量的缩并

一个张量 $\Phi_{\cdots kl \cdots}$ 与单位张量的**内积**定义为

$$\sum_{kl} \Phi_{\cdots kl \cdots} \delta_{kl} = \sum_{k} \Phi_{\cdots kk \cdots},$$

结果是一新张量，其阶数减少 2 阶，又称**张量的缩并**. 以 $\sum\limits_{k} \Phi_{ikk}$ 为例，显然有

$$\begin{aligned}\sum_{k} \overset{(y)}{\Phi}_{ikk} &= \sum_{l,m,n} \sum_{k} a_{il} a_{km} a_{kn} \overset{(x)}{\Phi}_{lmn} \\ &= \sum_{l,m,n} a_{il} \delta_{mn} \overset{(x)}{\Phi}_{lmn} \quad \left(\because \sum_{k} a_{km} a_{kn} = \delta_{mn}\right) \\ &= \sum_{l} a_{il} \sum_{m} \overset{(x)}{\Phi}_{lmm},\end{aligned}$$

表明经过缩并后的 $\sum\limits_{k} \Phi_{ikk}$ 为一阶张量即矢量.

5.3.6　张量的乘法

同阶或不同阶的两个张量可以相乘，其积仍为一张量，其阶数为两张量阶数之和. 这种乘法称为**张量乘法**，所得之积有时称为**直积**或**外积**. 例如，二阶张量 $\boldsymbol{\Phi}$ 与

三阶张量 Ψ 的张量积为五阶张量 Θ，记为

$$\Phi\Psi = \Theta \quad 或 \quad \Phi_{ij}\Psi_{lmn} = \Theta_{ijlmn}.$$

注意到

$$\begin{aligned}\overset{(y)}{\Theta}_{ijlmn} &= \overset{(y)}{\Phi}_{ij}\overset{(y)}{\Psi}_{lmn}\\ &= \sum_{h,k} a_{ih}a_{jk}\overset{(x)}{\Phi}_{hk}\sum_{p,q,r} a_{lp}a_{mq}a_{nr}\overset{(x)}{\Psi}_{pqr}\\ &= \sum_{h,k,p,q,r} a_{ih}a_{jk}a_{lp}a_{mq}a_{nr}(\overset{(x)}{\Phi}_{hk}\overset{(x)}{\Psi}_{pqr})\\ &= \sum_{h,k,p,q,r} a_{ih}a_{jk}a_{lp}a_{mq}a_{nr}\overset{(x)}{\Theta}_{hkpqr},\end{aligned}$$

即 Θ 是五阶张量.

5.3.7 张量的缩乘

同阶或不同阶的两个张量相乘时，若各有一指标相同并对之求和，则称为**张量的缩乘**；结果仍是一张量. 张量的缩乘积又称**张量的内积或标积**，其阶数为两张量阶数之和减 2.

例如，$\sum_j \Phi_{ij}\Psi_{jmn}$，$\sum_i \Phi_{ij}\Psi_{lmi}$，$\sum_j \Phi_{ij}\Psi_{ljn}$ 等缩乘积都是三阶张量；但是要注意到，由于被缩并指标位置的不同，它们是不同的张量.

一个二阶张量 Φ_{ij} 经过缩并后为一标量 $\sum_l \Phi_{ll}$. 例如，对于并矢 $\Phi = ab$，则有 $\sum_l \Phi_{ll} = \sum_l a_l b_l = \boldsymbol{a}\cdot\boldsymbol{b}$，称为矢量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的标积或内积. 下面再举几例：

例 1 $I = \boldsymbol{ii}+\boldsymbol{jj}+\boldsymbol{kk}$，

$$\begin{aligned}I\cdot\boldsymbol{r} &= \boldsymbol{ii}\cdot\boldsymbol{r}+\boldsymbol{jj}\cdot\boldsymbol{r}+\boldsymbol{kk}\cdot\boldsymbol{r}\\ &= \boldsymbol{i}x+\boldsymbol{j}y+\boldsymbol{k}z = \boldsymbol{r}.\end{aligned}$$

例 2 并矢 $\Phi = \boldsymbol{ab}$ 与矢量 $\boldsymbol{c}$ 的内积有

$$\Phi\cdot\boldsymbol{c} = \boldsymbol{ab}\cdot\boldsymbol{c} = \boldsymbol{a}\sum_l b_l c_l,$$

$$\boldsymbol{c}\cdot\Phi = \boldsymbol{c}\cdot\boldsymbol{ab} = \left(\sum_l c_l \boldsymbol{a}_l\right)\boldsymbol{b},$$

它们是不同的矢量.

例 3 设有并矢 $\Phi = \boldsymbol{ab}$ 和 $\Psi = \boldsymbol{cd}$，Φ 与 Ψ 作二次缩乘后有

$$\Phi : \Psi = \boldsymbol{ab} : \boldsymbol{cd} = \sum_{l,m} a_l b_m c_m d_l = \Psi : \Phi,$$

它是一个标量.

5.3.8 张量的导数

对于正交变换

$$y_i = \sum_r a_{ir} x_r,$$

由于 $\widetilde{A}A=1$ 或 $A^{-1}=\widetilde{A}$，所以有逆变换

$$x_r = \sum_i a_{ir} y_i.$$

因此有下列求导关系

$$\frac{\partial}{\partial y_i} = \sum_r \frac{\partial x_r}{\partial y_i}\frac{\partial}{\partial x_r} = \sum_r a_{ir}\frac{\partial}{\partial x_r}.$$

由此可见，求导符$\frac{\partial}{\partial x_r}$像矢量一样变换.

对于**标量场** $\overset{(y)}{\varphi}=\overset{(x)}{\varphi}$，容易求得

$$\frac{\partial}{\partial y_i}\overset{(y)}{\varphi} = \sum_r a_{ir}\frac{\partial}{\partial x_r}\overset{(x)}{\varphi},$$

即$\frac{\partial\varphi}{\partial x_r}$是一矢量. 它称为 φ 的**梯度**，一般记作 $\nabla\varphi$ 或 $\mathrm{grad}\varphi$.

对于**张量场** $\Phi_{pq\cdots}$，偏导数 $\partial\Phi_{pq\cdots}/\partial x_r$ 构成一张量，其阶数比 $\Phi_{pq\cdots}$ 的增加一阶. 现以二阶张量 Φ_{pq} 为例予以证明.

$$\frac{\partial}{\partial y_k}\overset{(y)}{\Phi}_{ij} = \frac{\partial}{\partial y_k}\Big(\sum_{p,q} a_{ip}a_{jq}\overset{(x)}{\Phi}_{pq}\Big) = \sum_r a_{kr}\frac{\partial}{\partial x_r}\Big(\sum_{p,q} a_{ip}a_{jq}\overset{(x)}{\Phi}_{pq}\Big)$$
$$= \sum_{pqr} a_{ip}a_{jq}a_{kr}\frac{\partial}{\partial x_r}\overset{(x)}{\Phi}_{pq};$$

所以，$\frac{\partial}{\partial x_r}\overset{(x)}{\Phi}_{pq}$是三阶张量.

另一方面，**矢量场** V_i 的偏导数 $\partial V_i/\partial x_j$ 是张量；它经缩并后 $\sum_i \partial V_i/\partial x_i$ 是标量，称为矢量 $\boldsymbol{V}$ 的**散度**，记作 $\nabla\cdot\boldsymbol{V}$ 或 $\mathrm{div}\boldsymbol{V}$，即

$$\mathrm{div}\boldsymbol{V} = \sum_i \frac{\partial V_i}{\partial x_i}.$$

类似地，对张量场 $\Phi_{pq\cdots}$ 求散度的结果：

$$\mathrm{div}_q\Phi_{pq\cdots} = \sum_q \frac{\partial\Phi_{pq\cdots}}{\partial x_q}$$

是张量，其阶数比 $\Phi_{pq\cdots}$ 减少一阶. 这里应该注意，所得张量会因被缩并指标不同而异.

现在引进缩写记号 $\partial V_i/\partial x_j = V_{i,j}$，即用逗号后的指标表明对相应坐标的求偏

导;因而 $V_{i,jk}$ 是二阶导数. 于是,前面的有关公式可分别写成

$$\overset{(y)}{\varphi}_{,i} = \sum_r a_{ir} \overset{(x)}{\varphi}_{,r},$$

$$\overset{(y)}{\Phi}_{ij,k} = \sum_{pqr} a_{ip} a_{jq} a_{kr} \overset{(x)}{\Phi}_{pq,r},$$

$$\mathrm{div} V = \sum_i V_{i,i},$$

$$\mathrm{div}_q \Phi_{pq\cdots} = \sum_q \Phi_{pq\cdots,q}.$$

5.3.9 张量方程

有了以上关于张量的各种代数运算法则,现在可以写出,例如,下列**张量方程**

$$\lambda \Phi \Psi + \mu \Theta = \mathbf{0}.$$

具体例子如

$$\boldsymbol{F} - m\boldsymbol{a} = \mathbf{0},$$

$$\mathsf{T} = (\boldsymbol{ED} + \boldsymbol{HB}) - \frac{1}{2}(\boldsymbol{E} \cdot \boldsymbol{D} + \boldsymbol{H} \cdot \boldsymbol{B})\mathsf{I}.$$

应该注意,一个方程中每一项的张量阶数应该相等.

5.4 对称和反对称张量

5.4.1 张量指标的置换

张量指标的置换构成另一张量.

例如:$\widetilde{\Phi}_{\cdots kl \cdots} = \Phi_{\cdots lk \cdots}$ 也是一张量. 利用 5.2.2 节定义的对称性即可证明. 因为

$$\overset{(y)}{\Phi}_{\cdots ij \cdots} = \sum_{\cdots k, l \cdots} \cdots a_{ik} a_{jl} \cdots \overset{(x)}{\Phi}_{\cdots kl \cdots},$$

当 j 与 i 对换后上式仍满足张量的定义条件,所以 $\widetilde{\Phi}_{\cdots ij \cdots}$ 也是张量.

又例如:$\widetilde{\Phi}_{\cdots klm \cdots} = \Phi_{\cdots lmk \cdots}$ 表示指标 klm 的一个轮换(klm),也是一张量. 因为

$$\overset{(y)}{\Phi}_{\cdots klm \cdots} = \sum_{\cdots pqr \cdots} \cdots a_{kp} a_{lq} a_{mr} \cdots \overset{(x)}{\Phi}_{\cdots pqr \cdots},$$

经轮换(klm)后上式仍满足张量的定义条件,所以 $\widetilde{\Phi}_{\cdots klm \cdots}$ 也是张量.

另一方面,如果在(x)坐标系内 $\overset{(x)}{\widetilde{\Phi}}_{\cdots ij \cdots} = \overset{(x)}{\Phi}_{\cdots ji \cdots}$,则有

$$\overset{(y)}{\widetilde{\Phi}}_{\cdots lk \cdots} = \sum_{\cdots ij \cdots} \cdots a_{li} a_{kj} \cdots \widetilde{\Phi}_{\cdots ij \cdots}$$

$$= \sum \cdots a_{kj} a_{li} \cdots \overset{(x)}{\Phi}_{\cdots ji \cdots} = \overset{(y)}{\Phi}_{\cdots kl \cdots};$$

也就是说，在任何坐标系都有

$$\widetilde{\Phi}_{\cdots lk \cdots} = \Phi_{\cdots kl \cdots},$$

对于 $\widetilde{\Phi}_{\cdots lk \cdots} = -\Phi_{\cdots kl \cdots}$ 的情况也是如此. 所以，有下一小节的定义.

5.4.2 对称和反对称张量

(1) 对称和反对称张量的定义

若张量 $\boldsymbol{\Phi}$(相对于某些指标的置换)具有性质

$$\widetilde{\boldsymbol{\Phi}} = \boldsymbol{\Phi} \quad 或 \quad \widetilde{\boldsymbol{\Phi}} = -\boldsymbol{\Phi},$$

则相应地称 $\boldsymbol{\Phi}$ 为(相对于这些指标的)**对称张量**或**反对称张量**.

按照 5.4.1 节，$\widetilde{\boldsymbol{\Phi}}$ 同样是张量，而按照 5.3.9 节可建立张量之间的方程，所以，上述关于对称张量和反对称张量的定义是恰当的.

这里顺便指出，就 ij 指标而言，对于对称张量

$$\Phi_{\cdots ij \cdots} = \Phi_{\cdots ji \cdots},$$

而对于反对称张量

$$\Phi_{\cdots ij \cdots} = -\Phi_{\cdots ji \cdots},$$

$$\Phi_{\cdots ii \cdots} = 0.$$

(2) 对称张量的例子

① 对称并矢(原型)：

$$\boldsymbol{\Phi} = \boldsymbol{r}\boldsymbol{r}, \qquad 即 \quad \Phi_{kl} = x_k x_l;$$

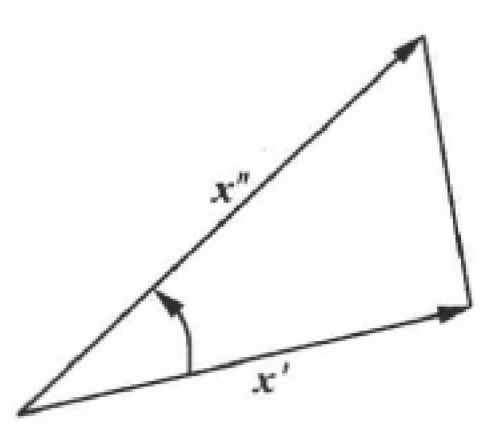
$$\boldsymbol{\Phi} = \boldsymbol{r}'\boldsymbol{r}'' + \boldsymbol{r}''\boldsymbol{r}', \quad 即 \quad \Phi_{kl} = x'_k x''_l + x''_k x'_l.$$

② 质点系的转动惯量张量：

$$I_{ij} = \sum_a m_a [r_a^2 \delta_{ij} - (\boldsymbol{r}_a)_i (\boldsymbol{r}_a)_j].$$

(3) 反对称张量的例子

① $2 \times n$ 矩阵 $\begin{bmatrix} x'_1 & x'_2 & \cdots & x'_n \\ x''_1 & x''_2 & \cdots & x''_n \end{bmatrix}$ 的所有 2 阶子式作为其分量所构成的张量 $\boldsymbol{\Phi}$：



$$\overset{(x)}{\boldsymbol{\Phi}} = \boldsymbol{x}'\boldsymbol{x}'' - \boldsymbol{x}''\boldsymbol{x}', 或$$

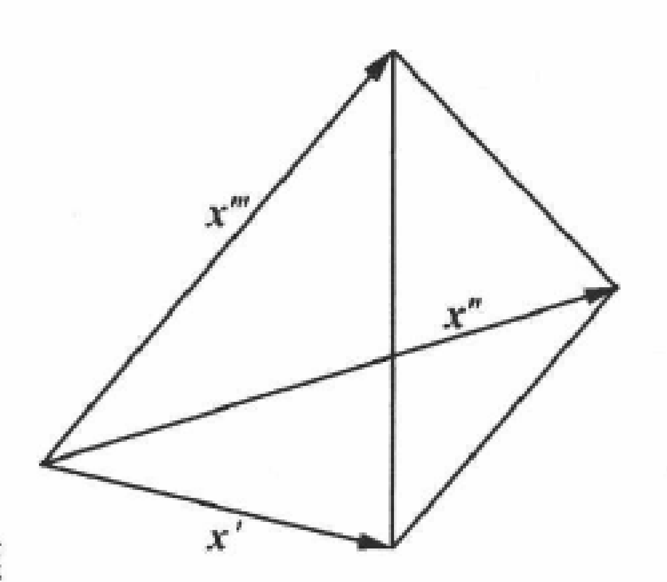
$$\overset{(x)}{\Phi}_{kl} = x'_k x''_l - x'_l x''_k = \begin{vmatrix} x'_k & x''_k \\ x'_l & x''_l \end{vmatrix},$$

称为**二重矢量**，是二阶反对称张量的原型. 特别是在三维 Euclid 空间中，它表示左图所示的定向面积.

② 类似地，

$$\overset{(x)}{\Phi}_{klm}=\begin{vmatrix} x_k' & x_k'' & x_k''' \\ x_l' & x_l'' & x_l''' \\ x_m' & x_m'' & x_m''' \end{vmatrix}$$

$$=\det[x_k',x_l'',x_m'''],$$



称为**三重矢量**，是三阶反对称张量的原型. 特别是在三维 Euclid 空间中，它表示右图所示的定向体积.

5.4.3 全反对称张量·赝张量

对所有指标中两个指标的对换均为反对称的张量称为**全反对称张量**.

(1) 赝标量

n 维 Euclid 空间中的 n 阶全反对称张量，其原型为 $\det[\boldsymbol{x}',\boldsymbol{x}'',\cdots,\boldsymbol{x}^{(n)}]$，于是，

$$\overset{(x)}{\Phi}_{p_1p_2\cdots p_n}=\det[x_{p_1}',x_{p_2}'',\cdots,x_{p_n}^{(n)}]$$

$$=\operatorname{sgn}\begin{pmatrix} p_1 & p_2 & \cdots & p_n \\ 1 & 2 & \cdots & n \end{pmatrix}\cdot\det[x_1'\ x_2''\cdots x_n^{(n)}],$$

或者更一般地，

$$\overset{(x)}{\Phi}_{p_1p_2\cdots p_n}=\overset{(x)}{\Psi}\operatorname{sgn}\begin{pmatrix} p_1 & p_2 & \cdots & p_n \\ 1 & 2 & \cdots & n \end{pmatrix},$$

这是**赝标量**.

首先用 3 维空间作为例证：

$$\overset{(y)}{\Phi}_{klm}=\det[y_k'\ y_l''\ y_m''']$$

$$=\operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ k & l & m \end{pmatrix}\cdot\det[y_1'\ y_2''\ y_3'''],$$

由

$$\begin{bmatrix} y_1' & y_1'' & y_1''' \\ y_2' & y_2'' & y_2''' \\ y_3' & y_3'' & y_3''' \end{bmatrix}=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1' & x_1'' & x_1''' \\ x_2' & x_2'' & x_2''' \\ x_3' & x_3'' & x_3''' \end{bmatrix}$$

得到

$$\det[y_1',y_2'',y_3''']=|A|\ \det[x_1',x_2'',x_3'''].$$

因为对于正交变换有 $|A|=\pm1$，所以

$$\overset{(y)}{\Phi}=\pm\overset{(x)}{\Phi}.$$

对于普通的标量 φ，总有

$$\overset{(y)}{\varphi} = \overset{(x)}{\varphi},$$

即 $\overset{(x)}{\varphi}$ 的前面永远是正号；而 Φ_{123} 有些像标量，其绝对值不变，但又有些不同，它可能变号；所以称它为**赝标量**.

因此，体积是赝标量.

现在再来证明一般情况. 对于

$$\overset{(x)}{\Phi}_{p_1\cdots p_n} = \overset{(x)}{\Psi}\operatorname{sgn}\begin{pmatrix} p_1 & \cdots & p_n \\ 1 & \cdots & n \end{pmatrix},$$

于是

$$\begin{aligned}\overset{(y)}{\Phi}_{i_1 i_2\cdots i_n} &= \sum_{p_1\cdots p_n} a_{i_1 p_1} a_{i_2 p_2}\cdots a_{i_n p_n}\overset{(x)}{\Phi}_{p_1 p_2\cdots p_n} \\ &= \sum_{n!\text{个置换}}\operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} a_{i_1 p_1}\cdots a_{i_n p_n}\cdot\overset{(x)}{\Psi} \\ &= \operatorname{sgn}\begin{pmatrix} i_1 & \cdots & i_n \\ 1 & \cdots & n \end{pmatrix} |A|\ \overset{(x)}{\Psi}\ \text{或}\ 0,\end{aligned}$$

后一结果是由行列式的定义和性质(第 3 章的 3.1 节和 3.2 节)得出的. 因此，

$$\overset{(y)}{\Phi}_{i_1\cdots i_n} = \overset{(y)}{\Psi}\operatorname{sgn}\begin{pmatrix} i_1 & \cdots & i_n \\ 1 & \cdots & n \end{pmatrix},$$

即

$$\overset{(y)}{\Psi} = |A|\ \overset{(x)}{\Psi},$$

由于 $|A| = \pm 1$，所以这是赝标量.

(2) 赝矢量

n 维 Euclid 空间中的 $n-1$ 阶全反对称张量，其原型是由 $n\times(n-1)$ 矩阵$[\boldsymbol{x}'\boldsymbol{x}''\cdots\boldsymbol{x}^{n-1}]$的 $n-1$ 阶子式作为其分量所构成的张量 $\boldsymbol{\Phi}$：

$$\begin{aligned}\overset{(x)}{\Phi}_{p_1 p_2\cdots p_{n-1}} &= \det[x'_{p_1}, x''_{p_2}, \cdots, x^{(n-1)}_{p_{n-1}}] \\ &= \operatorname{sgn}\begin{pmatrix} 1 & \cdots & n-1 & n \\ p_1 & \cdots & p_{n-1} & p_n \end{pmatrix}\operatorname{sgn}\begin{pmatrix} q_1 & \cdots & q_{n-1} & p_n \\ 1 & \cdots & n-1 & n \end{pmatrix} \\ &\quad\cdot\det[x'_{q_1}, \cdots, x^{(n-1)}_{q_{n-1}}],\end{aligned}$$

这里 $q_1 < q_2 < \cdots < q_{n-1}$. 或者更一般地，

$$\overset{(x)}{\Phi}_{p_1\cdots p_{n-1}} = \operatorname{sgn}\begin{pmatrix} 1 & \cdots & n \\ p_1 & \cdots & p_n \end{pmatrix}\overset{(x)}{\Psi}_{p_n},$$

这是**赝矢量**，因为矢量前带有正负号.

首先用三维空间作为例证：由

$$y'_i = \sum_k a_{ik} x'_k,$$

$$y'_i y''_j = \sum_{kl} a_{ik} a_{jl} x'_k x''_l,$$

$$y'_j y''_i = \sum_{pq} a_{jp} a_{iq} x'_p x''_q = \sum_{lk} a_{jl} a_{ik} x'_l x''_k,$$

得

$$y'_i y''_j - y'_j y''_i = \sum_{kl} a_{ik} a_{jl} (x'_k x''_l - x'_l x''_k);$$

因而

$$\overset{(y)}{\Phi}_{ij} = \frac{1}{2} \sum_{kl} (a_{ik} a_{jl} - a_{il} a_{jk}) \overset{(x)}{\Phi}_{kl}.$$

由上式显然看出，当 $i=j$ 或 $k=l$ 时，右边为零；所以可取 (ijn) 和 (klm) 为 (123) 的偶置换. 于是，令 $\overset{(y)}{\Phi}_{ij} = \overset{(y)}{\Psi}_n$，$\overset{(x)}{\Phi}_{kl} = \overset{(x)}{\Psi}_m$，并有

$$(a_{ik} a_{jl} - a_{il} a_{jk}) = A_{nm} \quad \text{及} \quad \frac{1}{2} \sum_{kl} \to \sum_m;$$

同时注意到

$$a_{nm} = (\widetilde{A})_{mn} = (A^{-1})_{mn} = \frac{A_{nm}}{\det A}, \text{即}$$

$$A_{nm} = (\det A) a_{nm};$$

最后得到

$$\overset{(y)}{\Psi}_n = \sum_m A_{nm} \overset{(x)}{\Psi}_m = (\det A) \sum_m a_{nm} \overset{(x)}{\Psi}_m.$$

由此可见，赝矢量与普通矢量之区别在于坐标变换时前面多了一个因子 det A. 赝矢量又称**轴矢量**.

例如，数学中两矢量的矢积和曲面的法线，物理学中的磁场强度、磁感应强度、磁化强度、角动量等均是赝矢量.

现在再来证明一般情况. 对于

$$\overset{(x)}{\Phi}_{p_1 \cdots p_{n-1}} = \overset{(x)}{\Psi}_{p_n} \operatorname{sgn}\begin{pmatrix} 1 & \cdots & n \\ p_1 & \cdots & p_n \end{pmatrix},$$

于是

$$\begin{aligned}
\overset{(y)}{\Phi}_{i_1 \cdots i_{n-1}} &= \sum_{p_1 \cdots p_{n-1}} a_{i_1 p_1} \cdots a_{i_{n-1} p_{n-1}} \overset{(x)}{\Phi}_{p_1 \cdots p_{n-1}} \\
&= \sum_{p_1 \cdots p_{n-1}} a_{i_1 p_1} \cdots a_{i_{n-1} p_{n-1}} \operatorname{sgn}\begin{pmatrix} 1 & \cdots & n \\ p_1 & \cdots & p_n \end{pmatrix} \overset{(x)}{\Psi}_{p_n} \\
&= \operatorname{sgn}\begin{pmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{pmatrix} \sum_{p_1 \cdots p_{n-1}} \operatorname{sgn}\begin{bmatrix} i_1 & \cdots & i_n \\ p_1 & \cdots & p_n \end{bmatrix} \\
&\quad \cdot a_{i_1 p_1} \cdots a_{i_{n-1} p_{n-1}} \overset{(x)}{\Psi}_{p_n}
\end{aligned}$$

$$= \operatorname{sgn}\begin{pmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{pmatrix} \sum_{p_n} A_{i_n p_n} \overset{(x)}{\Psi}_{p_n}$$

$$= (\det A) \operatorname{sgn}\begin{pmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{pmatrix} \sum_{p_n} a_{i_n p_n} \overset{(x)}{\Psi}_{p_n}$$

$$= \overset{(y)}{\Psi}_{i_n} \operatorname{sgn}\begin{pmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{pmatrix},$$

$$\overset{(y)}{\Psi}_k = (\det A) \sum_l a_{kl} \overset{(x)}{\Psi}_l .$$

(3) 赝张量

一般,n 维 Euclid 空间中的 $n-r$ 阶全反对称张量称为 r 阶**赝张量**.

r 阶赝张量 $\Psi_{i_1 \cdots i_r}$ 的变换规则是

$$\overset{(y)}{\Psi}_{i_1 \cdots i_r} = (\det A) \sum_{k_1 \cdots k_r} a_{i_1 k_1} \cdots a_{i_r k_r} \overset{(x)}{\Psi}_{k_1 \cdots k_r} .$$

下面介绍一个赝张量 $\varepsilon_{p_1 p_2 \cdots p_n}$,当 $p_1, p_2, \cdots, p_n$ 为 $1,2,\cdots,n$ 的偶置换时其值是 $+1$,为奇置换时其值是 -1,其余一切分量均为 0;即取

$$\overset{(x)}{\varepsilon}_{p_1 p_2 \cdots p_n} = \operatorname{sgn}\begin{pmatrix} p_1 & p_2 & \cdots & p_n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

容易证明

$$\overset{(y)}{\varepsilon}_{1,2,\cdots,n} = (\det A) \sum_{p_1 \cdots p_n} a_{1p_1} a_{2p_2} \cdots a_{np_n} \overset{(x)}{\varepsilon}_{p_1 p_2 \cdots p_n}$$

$$= (\det A) \sum_{p_1 \cdots p_n} \operatorname{sgn}\begin{pmatrix} p_1 & p_2 & \cdots & p_n \\ 1 & 2 & \cdots & n \end{pmatrix} a_{1p_1} a_{2p_2} \cdots a_{np_n}$$

$$= (\det A)^2 = 1.$$

因为 $\overset{(y)}{\varepsilon}_{k_1 k_2 \cdots k_n}$ 也是全反对称函数,其分量也是取值 $0, \pm 1$,所以,$\varepsilon_{p_1 p_2 \cdots p_n}$ 是在一切坐标系中均具有相同分量的赝张量,称为 **Levi-Civita 符号**或 **Levi-Civita 赝张量.**

利用 Levi-Civita 符号,对于 3 维 Euclid 空间,可使很多数学公式的书写得以简化. 例如:

① 二矢量的**矢积**可写成 $(\boldsymbol{u} \times \boldsymbol{v})_i = \sum_{jk} \varepsilon_{ijk} u_j v_k$;

② 矢量的**旋度**可写成 $(\nabla \times \boldsymbol{v})_i = \sum_{jk} \varepsilon_{ijk} v_{k,j}$;

③ 角动量分量之间的**对易关系**可写成

$$l_i l_j - l_j l_i = \mathrm{i}\hbar \varepsilon_{ijk} l_k ;$$

以上公式中的 i, j, k 代表 x, y, z 中的任一个.

(4) 关于赝张量运算的性质

关于赝张量，不难证明以下几点结果：

(i) 两个同阶赝张量之和或差仍为赝张量.

(ii) 一张量与一赝张量之积为赝张量.

(iii) 两个赝张量之积为张量.

(iv) 赝张量经缩并后仍为赝张量.

(v) 赝张量对 x_i 的偏导数是赝张量.

5.5 赝 Euclid 张量

如果常系数线性变换

$$y_i = \sum_{k=0}^{3} a_{ik} x_k \quad (i = 0,1,2,3)$$

保持

$$-x_0^2 + x_1^2 + x_2^2 + x_3^2 = -y_0^2 + y_1^2 + y_2^2 + y_3^2$$

为不变量，这样的变换可称为**赝正交变换**. 若令 $x_4 = \mathrm{i}x_0$, $y_4 = \mathrm{i}y_0$，则化至以前考虑过的情况，

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = y_1^2 + y_2^2 + y_3^2 + y_4^2;$$

因为有一个虚坐标 x_4，所以(x_1, x_2, x_3, x_4)的空间称为**赝 Euclid 空间**，而相应线性变换

$$y_\mu = \sum_{\nu=1}^{4} a_{\mu\nu} x_\nu \quad (\mu = 1,2,3,4)$$

可认为是赝 Euclid 空间中的赝正交变换（系数可取复数值）. 应该注意到，**Minkowski 四维时空**（$x_4 = \mathrm{i}x_0 = \mathrm{i}ct$）中的 **Lorentz 变换**就是这里所指的最主要情况.

在赝 Euclid 空间中的张量，称为**赝 Euclid 张量**，其分量凡含有偶数(包括 0)个指标 4 的为实数而含奇数个指标 4 的为纯虚数. 例如，对于二阶赝 Euclid 张量 $\boldsymbol{\Phi}$：

$$\boldsymbol{\Phi} = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} \\ \Phi_{21} & \Phi_{22} & \Phi_{23} & \Phi_{24} \\ \Phi_{31} & \Phi_{32} & \Phi_{33} & \Phi_{34} \\ \Phi_{41} & \Phi_{42} & \Phi_{43} & \Phi_{44} \end{bmatrix},$$

其中 $\Phi_{14}, \Phi_{24}, \Phi_{34}, \Phi_{41}, \Phi_{42}, \Phi_{43}$ 是虚数，其余(包括 Φ_{44})是实数.

在相对论中，采用 Minkowski 四维时空：

$$(x_0, x_1, x_2, x_3) = (ct, x, y, z) \quad 或 \quad (x_1, x_2, x_3, x_4) = (x, y, z, \mathrm{i}ct).$$

四维矢量是四维时空中的矢量. 例如：

(1) 四维势矢量

$$(A_0,A_1,A_2,A_3)=\left(\frac{\varphi}{c},A_x,A_y,A_z\right)\quad 或$$

$$(A_1,A_2,A_3,A_4)=(A_x,A_y,A_z,\mathrm{i}\varphi/c),$$

其中 $\boldsymbol{A}=(A_x,A_y,A_z)$是矢势,$\varphi$ 是标势,c 是真空中光速.

(2) 四维流密度矢量

$$(j_0,j_1,j_2,j_3)=(\rho c,j_x,j_y,j_z)\quad 或$$

$$(j_1,j_2,j_3,j_4)=(j_x,j_y,j_z,\mathrm{i}\rho c),$$

其中 ρ 是电荷密度,$\boldsymbol{j}=(j_x,j_y,j_z)$是电流密度矢量.

所谓**6 矢量**就是四维时空中的二阶反对称张量,它的 6 个不为 0 的独立分量分别表示电场强度 $\boldsymbol{E}$ 和磁感应强度 $\boldsymbol{B}$. 写成张量形式为

$$(F_{ij})=\begin{bmatrix}F_{00}&F_{01}&F_{02}&F_{03}\\F_{10}&F_{11}&F_{12}&F_{13}\\F_{20}&F_{21}&F_{22}&F_{23}\\F_{30}&F_{31}&F_{32}&F_{33}\end{bmatrix}$$

$$=\begin{bmatrix}0&-\frac{1}{c}E_x&-\frac{1}{c}E_y&-\frac{1}{c}E_z\\\frac{1}{c}E_x&0&B_z&-B_y\\\frac{1}{c}E_y&-B_z&0&B_x\\\frac{1}{c}E_z&B_y&-B_x&0\end{bmatrix};$$

其中 $F_{ij}=\dfrac{\partial A_j}{\partial x_i}-\dfrac{\partial A_i}{\partial x_j}=-F_{ji}\quad(i,j=1,2,3)$,$F_{i0}=-\dfrac{\partial A_0}{\partial x_i}-\dfrac{\partial A_i}{\partial x_0}=-F_{0i}\quad(i=1,2,3)$,$F_{00}=0$;或者,

$$(F_{\mu\nu})=\begin{bmatrix}F_{11}&F_{12}&F_{13}&F_{14}\\F_{21}&F_{22}&F_{23}&F_{24}\\F_{31}&F_{32}&F_{33}&F_{34}\\F_{41}&F_{42}&F_{43}&F_{44}\end{bmatrix}$$

$$=\begin{bmatrix}0&B_z&-B_y&-\mathrm{i}E_x/c\\-B_z&0&B_x&-\mathrm{i}E_y/c\\B_y&-B_x&0&-\mathrm{i}E_z/c\\\mathrm{i}E_x/c&\mathrm{i}E_y/c&\mathrm{i}E_z/c&0\end{bmatrix};$$

其中 $F_{\mu\nu}=\dfrac{\partial A_\nu}{\partial x_\mu}-\dfrac{\partial A_\mu}{\partial x_\nu}=-F_{\nu\mu}\quad(\mu,\nu=1,2,3,4)$.

5.6 广义坐标变换下的张量①②

迄今所讨论的 Descartes 张量，是在 Euclid 或赝 Euclid 空间中直角坐标系的正交变换下所定义的张量；因为系数 a_{ik} 与坐标 $x_1,\cdots,x_n$ 无关，所以，不同点的矢量可予以相加或相减，称为**自由矢量**. 但是，在广义坐标变换下，情况不再是如此，因为这里的变换系数逐点变化，所以，只有隶属于同一场点的矢量或张量才能进行代数运算.

5.6.1 广义坐标变换

现在考虑广义坐标变换. 为了需要和方便，首先引进新的记号和约定.

设 $x^1,x^2,\cdots,x^n$ 和 $x'^1,x'^2,\cdots,x'^n$ 为 $\mathbf{V}_n(\mathbf{R})$空间中 P 点的两组坐标，以后分别称为旧坐标和新坐标(注意，这里 $1,2,\cdots,n$ 仅为上标，非指其幂)，它们有下列函数关系：

$$x^\mu \equiv x^\mu(x'^1,\cdots,x'^n) \quad (1\leqslant\mu\leqslant n),$$
$$x'^\alpha \equiv x'^\alpha(x^1,\cdots,x^n) \quad (1\leqslant\alpha\leqslant n);$$

称为**广义坐标变换**或**双连续变换**. 这里及以后常将 $1\leqslant\mu\leqslant n, 1\leqslant\alpha\leqslant n$ 之类条件略去. 对上面二式微分得到

$$\mathrm{d}x^\mu = \sum_\alpha \frac{\partial x^\mu}{\partial x'^\alpha}\mathrm{d}x'^\alpha = \frac{\partial x^\mu}{\partial x'^\alpha}\mathrm{d}x'^\alpha,$$
$$\mathrm{d}x'^\alpha = \sum_\mu \frac{\partial x'^\alpha}{\partial x^\mu}\mathrm{d}x^\mu = \frac{\partial x'^\alpha}{\partial x^\mu}\mathrm{d}x^\mu,$$

这里后面等式采用了 Einstein 求和约定. 一项中重复的指标(称为**傀标**，可换字母)隐含对该指标的所有可能取值(从 1 至 n)的求和；相反，一项中仅出现一次的指标(称为**自由指标**)具有 1 至 n 的任一确定值.

应该注意到，广义坐标变换中 x^μ 与 x'^α 之间的关系一般为非线性关系，但在某一场点邻域的微分关系则是线性的. 再一点是，这里隐含下列 Jacobi 行列式不为零：

$$\frac{\partial(x^1,\cdots,x^n)}{\partial(x'^1,\cdots,x'^n)} = \det\left\{\frac{\partial x^\mu}{\partial x'^\alpha}\right\} \neq 0,$$
$$\frac{\partial(x'^1,\cdots,x'^n)}{\partial(x^1,\cdots,x^n)} = \det\left\{\frac{\partial x'^\alpha}{\partial x^\mu}\right\} \neq 0,$$

---

① 彭桓武，徐锡申：《理论物理基础》，北京大学出版社，北京，1998(2011 年重排). 该书第 14 章习题 14.1—14.7 有简明扼要的介绍.

② 参考：A. W. Joshi, *Matrices and Tensors in Physics*, Wiley, New York, 2nd ed., 1984.

即坐标变换在该场点是非奇异变换，因而可逆.

同时，由于 $x^1,\cdots,x^n$ 是相互无关的，$x'^1,\cdots,x'^n$ 也是相互无关的，所以

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}x^\nu}=\delta^\mu{}_\nu=\begin{cases}1 & 若\ \mu=\nu,\\ 0 & 若\ \mu\neq\nu;\end{cases}$$

$$\frac{\mathrm{d}x'^\alpha}{\mathrm{d}x'^\beta}=\delta^\alpha{}_\beta=\begin{cases}1 & 若\ \alpha=\beta,\\ 0 & 若\ \alpha\neq\beta;\end{cases}$$

这里 $\delta^\mu{}_\nu,\delta^\alpha{}_\beta$ 是 Kronecker δ 符号. 注意到，分子上的指标作为上标，分母上的指标作为下标；后面将会阐明这样书写的理由.

另一方面，若将 x^μ 看作 x'^α 的函数，则有

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}x^\nu}=\frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x'^\alpha}{\partial x^\nu}=\delta^\mu{}_\nu,$$

按求和约定，中间项隐含对 α 求和；类似还有

$$\frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x^\mu}{\partial x'^\beta}=\delta^\alpha{}_\beta.$$

现在可以认为(采用求和约定)：

$$\mathrm{d}x'^\alpha=\frac{\partial x'^\alpha}{\partial x^\mu}\mathrm{d}x^\mu\quad 相当于\quad y_i=a_{ik}x_k,$$

即 $\mathrm{d}x'^\alpha$ 相当于 y_i，$\mathrm{d}x^\mu$ 相当于 x_k，$\dfrac{\partial x'^\alpha}{\partial x^\mu}$ 相当于 a_{ik}；同时 $\det\left\{\dfrac{\partial x'^\alpha}{\partial x^\mu}\right\}\neq 0$ 相当于 $\det\{a_{ik}\}\neq 0$，即上述坐标变换都是非奇异线性变换.

于是，对广义坐标变换，可依据同一场点处其新旧坐标系中分量的变换性质来定义各类型张量. 同时注意到，新分量是该场点新坐标的函数(全为带撇号“′”的量)，旧分量是该同一场点旧坐标的函数(全为不带撇号的量)，这种函数关系一般不予明显标出. 应该指出，这类张量是广义相对论中的主要数学工具.

5.6.2　反变矢量

广义坐标变换下，凡其分量具有下列变换性质者：

$$A'^\alpha=\frac{\partial x'^\alpha}{\partial x^\mu}A^\mu,$$

称为**反变一阶张量**，简称**反变矢量**.

若以 $\dfrac{\partial x^\nu}{\partial x'^\alpha}$ 乘上式并对 α 求和，注意到 $\dfrac{\partial x^\nu}{\partial x'^\alpha}\dfrac{\partial x'^\alpha}{\partial x^\mu}=\delta^\nu{}_\mu$，于是可得出其逆变换式

$$\frac{\partial x^\nu}{\partial x'^\alpha}A'^\alpha=\frac{\partial x^\nu}{\partial x'^\alpha}\frac{\partial x'^\alpha}{\partial x^\mu}A^\mu=\delta^\nu{}_\mu A^\mu=A^\nu,$$

即

$$A^\nu=\frac{\partial x^\nu}{\partial x'^\alpha}A'^\alpha.$$

显然，由

$$\mathrm{d}x'^{\alpha} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}}\mathrm{d}x^{\mu}$$

可知，位移 $\mathrm{d}\boldsymbol{x}=\{\mathrm{d}x^1,\cdots,\mathrm{d}x^n\}$是反变矢量的原型. 对于固定坐标系，在其两边同除以 $\mathrm{d}t$(t 表示时间)得速度的分量满足

$$v'^{\alpha} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}}v^{\mu},$$

即，速度也是反变矢量. 再求一次对时间的导数，并注意到对固定坐标系而言，系数 $\frac{\partial x'^{\alpha}}{\partial x^{\mu}}$不依赖于时间，故得加速度也是反变矢量. 另一方面，应该注意到，x^{μ} 并非张量的分量.

5.6.3 标量场

广义坐标变换下，若函数 $\varphi(x^1,\cdots,x^n)=\varphi(P)$具有下列变换性质者(即函数形式可能变化，但其在 P 点的数值始终不变者)：

$$\varphi'(x'^1,\cdots,x'^n) = \varphi(x^1,\cdots,x^n),$$

称为零阶张量. 简称**标量**.

若在 $\mathbf{V}_n(\mathbf{R})$空间每一点都附有一标量，则 $\varphi\equiv\varphi(x^{\mu})$的标量总体就形成一**标量场**.

5.6.4 协变矢量

广义坐标变换下，凡其分量具有下列变换性质者：

$$A'_{\alpha} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}A_{\mu},$$

称为**协变一阶张量**，简称**协变矢量**.

同样可求得其逆变换式为

$$A_{\nu} = \frac{\partial x'^{\alpha}}{\partial x^{\nu}}A'_{\alpha}.$$

可以证明，**标量场的梯度**是协变矢量. 因为，由 $\varphi'(x'^{\alpha})=\varphi(x^{\mu}(x'^{\alpha}))$ 容易求得

$$\frac{\partial \varphi'}{\partial x'^{\alpha}} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial \varphi}{\partial x^{\mu}},$$

若令 $A'_{\alpha}=\frac{\partial \varphi'}{\partial x'^{\alpha}}$，$A_{\mu}=\frac{\partial \varphi}{\partial x^{\mu}}$，它们的确具有协变矢量的变换性质. 实际上，标量场在 P 点的梯度是协变矢量的原型.

现在看到，$\frac{\partial \varphi}{\partial x^{\mu}}$与 $\mathrm{d}x^{\mu}$ 是变换性质不同的两类矢量. 前者由变换 $x^{\mu}(x'^{\alpha})$即

$x^\mu \to x'^\alpha, \varphi \to \varphi', \dfrac{\partial \varphi}{\partial x^\mu} \to \dfrac{\partial \varphi'}{\partial x'^\alpha}$而来，故称为协变；且矢量由与恒定坐标曲面正交的方向上的分量描述，相应 ∂x^μ 出现于分母，故用下标表示. 后者由逆变换 $x'^\alpha(x^\mu)$ 直接求导而来，故称为反变；且矢量由坐标增加方向上的分量描述，相应 dx^μ 出现于分子，故用上标表示. 注意到 Descartes 直角坐标系中，坐标方向 x^i 与恒定 x^i 面的正交方向重合，故协变与反变矢量之间的区别消失.

为了进一步对比，采用下列简写：

$$a^\alpha{}_\mu \equiv \frac{\partial x'^\alpha}{\partial x^\mu}, \quad b^\mu{}_\alpha \equiv \frac{\partial x^\mu}{\partial x'^\alpha},$$

$$\mathscr{A} = \{a^\alpha{}_\mu\}, \quad \mathscr{B} = \{b^\mu{}_\alpha\}.$$

前面已经求得

$$\sum_\mu a^\alpha{}_\mu b^\mu{}_\beta = a^\alpha{}_\mu b^\mu{}_\beta = \delta^\alpha{}_\beta \quad 即 \quad \mathscr{B} = \mathscr{A}^{-1}.$$

对于 Euclid 空间中的正交变换，$B = A^{-1} = \widetilde{A}$，即 $b_{ik} = a_{ki}$. 所以，在 Euclid 空间中，反变和协变矢量或张量之间没有任何区别.

5.6.5　混变张量

广义坐标变换下，凡其分量具有下列变换性质者：

$$T'^{\alpha_1 \cdots \alpha_r}_{\beta_1 \cdots \beta_s} = \frac{\partial x'^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial x'^{\alpha_r}}{\partial x^{\mu_r}} \frac{\partial x^{\nu_1}}{\partial x'^{\beta_1}} \cdots \frac{\partial x^{\nu_s}}{\partial x'^{\beta_s}} T^{\mu_1 \cdots \mu_r}_{\nu_1 \cdots \nu_s},$$

称为反变 r 阶协变 s 阶的**混变**$(r+s)$阶**张量**，简称(r,s)型张量. 例如：反变二阶张量(能量动量密度张量)

$$T'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} T^{\mu\nu},$$

协变二阶张量(电磁场张量)

$$F'_{\alpha\beta} = \frac{\partial x^\mu}{\partial x'^\alpha} \frac{\partial x^\nu}{\partial x'^\beta} F_{\mu\nu},$$

反变一阶协变一阶的混变二阶张量[(1,1)型张量]

$$N'^{\beta}_{\alpha} = \frac{\partial x^\mu}{\partial x'^\alpha} \frac{\partial x'^\beta}{\partial x^\nu} N_\mu{}^\nu,$$

反变一阶协变三阶的混变四阶张量[(1,3)型张量]

$$R'^{\epsilon}{}_{\alpha\beta\gamma} = \frac{\partial x'^\epsilon}{\partial x^\rho} \frac{\partial x^\lambda}{\partial x'^\alpha} \frac{\partial x^\mu}{\partial x'^\beta} \frac{\partial x^\nu}{\partial x'^\gamma} R^\rho{}_{\lambda\mu\nu}.$$

可以看出，前面引进的 Kronecker δ 符号 $\delta^\mu{}_\nu$ 是反变一阶协变一阶的混变二阶张量，因为

$$\delta'^\alpha{}_\beta = \frac{\partial x'^\alpha}{\partial x'^\beta} = \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x^\mu}{\partial x'^\beta} = \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x^\nu}{\partial x'^\beta} \frac{\partial x^\mu}{\partial x^\nu} = \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x^\nu}{\partial x'^\beta} \delta^\mu{}_\nu.$$

同时，这也说明了前面用上下指标书写的理由.

5.7 混变张量的代数运算[①]

下面考虑在空间同一点混变张量的代数运算.

5.7.1 张量的加法和减法

若两张量具有相同反变阶数和相同协变阶数，则称它们为**相同类型的**.

相同类型的两个张量可以相加或相减，其和或差与原张量为相同类型张量，其分量为两张量相应分量的和与差. 例如，

$$C^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s} = A^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s} + B^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s},$$

$$D^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s} = A^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s} - B^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s}.$$

两个张量，当且仅当它们为相同类型，且其所有对应分量均相等时，则称此**两张量相等**，即

$$A^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s} = B^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s}.$$

各类型张量都有全部分量均为零的**零张量**. 显然，若一张量在任一坐标系中的全部分量为零，则它们在一切坐标系中均为零.

5.7.2 张量的缩并

对一个混变张量，若令其一对上下指标相等(即对之求和)，称为**张量的缩并**；缩并一次，张量的反变阶数和协变阶数各减少一阶. 例如，(3,2)型张量(即反变三阶协变二阶的混变五阶张量)$A^{\lambda\mu\nu}_{\rho\sigma}$中若令$\rho=\lambda$，则缩并为

$$A^{\lambda\mu\nu}_{\lambda\sigma} = \sum_{\lambda} A^{\lambda\mu\nu}_{\lambda\sigma} = C^{\mu\nu}_{\sigma},$$

它是(2,1)型张量. 的确，

$$\begin{aligned} C'^{\beta\gamma}_{\eta} = A'^{\alpha\beta\gamma}_{\alpha\eta} &= \frac{\partial x'^{\alpha}}{\partial x^{\lambda}} \frac{\partial x'^{\beta}}{\partial x^{\mu}} \frac{\partial x'^{\gamma}}{\partial x^{\nu}} \frac{\partial x^{\rho}}{\partial x'^{\alpha}} \frac{\partial x^{\sigma}}{\partial x'^{\eta}} A^{\lambda\mu\nu}_{\rho\sigma} \\ &= \frac{\partial x'^{\beta}}{\partial x^{\mu}} \frac{\partial x'^{\gamma}}{\partial x^{\nu}} \frac{\partial x^{\sigma}}{\partial x'^{\eta}} \delta_{\lambda}{}^{\rho} A^{\lambda\mu\nu}_{\rho\sigma} \\ &= \frac{\partial x'^{\beta}}{\partial x^{\mu}} \frac{\partial x'^{\gamma}}{\partial x^{\nu}} \frac{\partial x^{\sigma}}{\partial x'^{\eta}} A^{\lambda\mu\nu}_{\lambda\sigma} \\ &= \frac{\partial x'^{\beta}}{\partial x^{\mu}} \frac{\partial x'^{\gamma}}{\partial x^{\nu}} \frac{\partial x^{\sigma}}{\partial x'^{\eta}} C^{\mu\nu}_{\sigma}, \end{aligned}$$

---

① 参考：A. W. Joshi, *Matrices and Tensors in Physics*.

这里用了$\frac{\partial x'^{\alpha}}{\partial x^{\lambda}}\frac{\partial x^{\rho}}{\partial x'^{\alpha}}=\delta_{\lambda}{}^{\rho}$.

显然，张量缩并可有多种方式；例如，$A^{\lambda\mu\nu}_{\rho\lambda}$，$A^{\lambda\mu\nu}_{\mu\sigma}$，$A^{\lambda\mu\nu}_{\rho\nu}$等都是张量$A^{\lambda\mu\nu}_{\rho\sigma}$的缩并形式，它们是不同张量．同时，缩并运算可重复进行，上述张量再次缩并后可得反变一阶张量$A^{\lambda\mu\nu}_{\lambda\mu}$或$A^{\lambda\mu\nu}_{\lambda\nu}$.

另一方面，缩并只能在上下指标之间进行．若令一张量的两上标或两下标相等，结果并非张量；例如，$A^{\lambda\lambda\nu}_{\rho\sigma}$和$A^{\lambda\mu\nu}_{\rho\rho}$并非张量.

5.7.3 张量的乘法

任何类型的两个张量可以相乘，乘积张量的反变阶数和协变阶数为因子张量相应阶数之和.

例如，$A^{\lambda\mu}_{\rho}B^{\nu}_{\sigma}=C^{\lambda\mu\nu}_{\rho\sigma}$仍为一张量，因为

$$C'^{\alpha\beta\gamma}_{\xi\eta}=A'^{\alpha\beta}_{\xi}B'^{\gamma}_{\eta}=\frac{\partial x'^{\alpha}}{\partial x^{\lambda}}\frac{\partial x'^{\beta}}{\partial x^{\mu}}\frac{\partial x^{\rho}}{\partial x'^{\xi}}A^{\lambda\mu}_{\rho}\frac{\partial x'^{\gamma}}{\partial x^{\nu}}\frac{\partial x^{\sigma}}{\partial x'^{\eta}}B^{\nu}_{\sigma}$$

$$=\frac{\partial x'^{\alpha}}{\partial x^{\lambda}}\frac{\partial x'^{\beta}}{\partial x^{\mu}}\frac{\partial x'^{\gamma}}{\partial x^{\nu}}\frac{\partial x^{\rho}}{\partial x'^{\xi}}\frac{\partial x^{\sigma}}{\partial x'^{\eta}}C^{\lambda\mu\nu}_{\rho\sigma}.$$

这样的乘积称为**张量的外积**或 **Kronecker 积**，简称**张量积**.

张量相乘后再缩并简称**缩乘**．例如，张量$A^{\lambda\mu}_{\rho}$和B^{ν}_{σ}可缩乘为$A^{\lambda\mu}_{\rho}B^{\rho}_{\sigma}$，$A^{\lambda\mu}_{\rho}B^{\nu}_{\mu}$，$A^{\lambda\mu}_{\rho}B^{\nu}_{\lambda}$，它们都是混变三阶张量；还可进一步缩乘为$A^{\lambda\mu}_{\rho}B^{\rho}_{\lambda}$，$A^{\lambda\mu}_{\rho}B^{\rho}_{\mu}$，它们都是反变一阶张量．张量的缩乘积又称**张量的内积**.

5.7.4 对称和反对称张量

张量指标排列有序，可以具有内禀对称性；与 Descartes 张量有类似之处．但这里要指出的是，对于广义坐标变换下的张量，只有对同类指标才能定义其对称或反对称．例如，若$A^{\lambda\mu\nu}_{\rho\sigma\tau}=A^{\mu\lambda\nu}_{\rho\sigma\tau}$，它对头两个反变指标为对称；或者，若$A^{\lambda\mu\nu}_{\rho\sigma\tau}=A^{\lambda\mu\nu}_{\tau\sigma\rho}$，它对第一和第三这对协变指标为对称，如此等等．对反对称亦可类似地定义，这种对称性在坐标变换下保持不变.

但是，对不同类型指标则不能如此定义．例如，若在x^{μ}坐标系有$A^{\lambda}_{\rho}=A^{\rho}_{\lambda}$，因为

$$A'^{\alpha}_{\beta}=\frac{\partial x'^{\alpha}}{\partial x^{\lambda}}\frac{\partial x^{\rho}}{\partial x'^{\beta}}A^{\lambda}_{\rho},$$

$$A'^{\beta}_{\alpha}=\frac{\partial x'^{\beta}}{\partial x^{\rho}}\frac{\partial x^{\lambda}}{\partial x'^{\alpha}}A^{\rho}_{\lambda}=\frac{\partial x'^{\beta}}{\partial x^{\rho}}\frac{\partial x^{\lambda}}{\partial x'^{\alpha}}A^{\lambda}_{\rho},$$

两式右边显然不相等，故一般而言，$A'^{\alpha}_{\beta}\neq A'^{\beta}_{\alpha}$．也就是说，一个张量在对不同类型指标之间的对称或反对称是某个坐标系下的偶然性质而非其内禀性质.

5.8 度规张量[①②]

5.8.1 度规张量

令 x^μ 和 $x^\mu+\mathrm{d}x^\mu$ 为 n 维空间中两邻近点相对于某个坐标系的坐标，两点之间的距离平方可表达为 $\mathrm{d}x^\mu$ 的**二次型**(参见第 6 章 6.1.1 节)

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu,$$

其中 $g_{\mu\nu}$ 可以是 x^μ 的函数(不失一般性，取 $g_{\mu\nu}=g_{\nu\mu}$)，且系数矩阵为非奇异矩阵[即 $g\equiv\det\{g_{\mu\nu}\}\neq 0$]. 可以证明，这样的 $g_{\mu\nu}$ 是二阶对称协变张量，称为**度规张量**.

因为 $\mathrm{d}s^2$ 是不变量，$\mathrm{d}x^\mu$ 是反变位移矢量的分量，在广义坐标变换 $x^\mu=x^\mu(x'^\alpha)$ 下，有

$$g'_{\alpha\beta}\mathrm{d}x'^\alpha\mathrm{d}x'^\beta = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = g_{\mu\nu}\frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta}\mathrm{d}x'^\alpha\mathrm{d}x'^\beta,$$

$$g'_{\alpha\beta} = \frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta}g_{\mu\nu}.$$

注意，具有这种一般性质的空间称为 **Riemann 空间**. 特别是，若系数 $g_{\mu\nu}$ 全与 x^μ 无关，则该空间变为 **Euclid 空间**. 当然，在 Euclid 空间中取曲线坐标系时也会有 $\mathrm{d}s^2=g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$ 的形式，但其函数 $g_{\mu\nu}$ 需满足某些条件. 还可考虑假设的**广义空间**，有可能在其中建立坐标系而并不假定其具有度规，甚至两点间距离也未予定义.

在 n 维空间中，对一个坐标系若有 $g_{\mu\nu}=0(\mu\neq\nu)$，则称为**正交坐标系**. 同时，若 $g_{\mu\nu}=\delta_{\mu\nu}(\delta_{\mu\nu}=1$，若 $\mu=\nu$，$\delta_{\mu\nu}=0$，若 $\mu\neq\nu$；$1\leqslant\mu,\nu\leqslant n)$，则称为 **Descartes 直角坐标系**.

5.8.2 反变度规张量

由如下关系

$$g_{\mu\nu}g^{\sigma\nu} = \delta_\mu{}^\sigma = \begin{cases}1 & (\mu=\sigma),\\ 0 & (\mu\neq\sigma)\end{cases}$$

定义一个与 $g_{\mu\nu}$ 共轭的反变张量 $g^{\mu\nu}$，称为**反变度规张量**. 注意到矩阵 $\{g^{\mu\nu}\}$ 正好是矩阵 $\{g_{\mu\nu}\}$ 之逆，有

$$g^{\mu\nu} = g^{\nu\mu} = (g_{\mu\nu}\text{ 的代数余子式})/g.$$

因此的确可以认为 $g_{\mu\nu}$ 和 $g^{\mu\nu}$ 分别为同一张量的协变和反变分量.

---

① 参考：A. W. Joshi, *Matrices and Tensors in Physics*.

② 参考：D. F. Lawden, *An Introduction to Tensor Calculus, Relativity and Cosmology*, 3rd ed., Wiley, New York, 1982.

5.8.3　相伴张量

任一反变矢量 A^{μ} 与协变度规张量 $g_{\mu\nu}$ 之缩乘

$$A^{\mu}g_{\mu\nu} = A_{\nu}$$

为一协变矢量 A_{ν}，可以认为 A^{μ} 与 A_{ν} 为同一矢量的反变与协变分量. 另外，

$$A_{\nu}g^{\nu\rho} = A^{\mu}g_{\mu\nu}g^{\nu\rho} = A^{\mu}\delta_{\mu}{}^{\rho} = A^{\rho},$$

表明 A_{ν} 与 A^{μ} 之间的关系可倒易. 这样的两个张量称为**相伴张量**.

5.8.4　指标的升降

利用度规张量 $g^{\mu\nu}$ 或 $g_{\mu\nu}$，可随意地实施**张量指标的提升或下降**. 例如，

$$A^{\mu\nu}g_{\nu\sigma} = A^{\mu}_{\cdot\sigma},\quad A^{\mu\nu}g_{\mu\rho} = A_{\rho\cdot}^{\cdot\nu},\quad A^{\mu\nu}g_{\mu\rho}g_{\nu\sigma} = A_{\rho\sigma};$$

$$A_{\rho\sigma}g^{\nu\sigma} = A_{\rho\cdot}^{\cdot\nu},\quad A_{\rho\cdot}^{\cdot\nu}g^{\mu\rho} = A^{\mu\nu},\quad A_{\rho\sigma}g^{\mu\rho}g^{\nu\sigma} = A^{\mu\nu};$$

$$A_{\lambda\cdot\cdot\cdot}^{\cdot\mu\nu\rho} = A^{\sigma\mu\nu\rho}g_{\sigma\lambda},\quad A^{\mu\cdot\nu\rho}_{\cdot\lambda\cdot\cdot} = A^{\mu\sigma\nu\rho}g_{\sigma\lambda},\quad \cdots,$$

$$A_{\lambda\mu\cdot\cdot}^{\cdot\cdot\nu\rho} = A^{\sigma\tau\nu\rho}g_{\sigma\lambda}g_{\tau\mu},\quad A^{\nu\cdot\cdot\rho}_{\cdot\lambda\mu\cdot} = A^{\nu\sigma\tau\rho}g_{\sigma\lambda}g_{\tau\mu},\quad \cdots,$$

$$A_{\lambda\mu\nu\rho} = A^{\sigma\tau\varphi\omega}g_{\sigma\lambda}g_{\tau\mu}g_{\varphi\nu}g_{\omega\rho};\quad \cdots.$$

首先注意到，$A^{\mu\nu}$，$A^{\mu}_{\cdot\nu}$，$A_{\mu\cdot}^{\cdot\nu}$，$A_{\mu\nu}$ 彼此为相伴张量，而 $A^{\lambda\mu\nu\rho}$，$A_{\lambda\cdot\cdot\cdot}^{\cdot\mu\nu\rho}$，$A^{\lambda\cdot\nu\rho}_{\cdot\mu\cdot\cdot}$ 等(4 个)，$A_{\lambda\mu\cdot\cdot}^{\cdot\cdot\nu\rho}$，$A^{\nu\cdot\cdot\rho}_{\cdot\lambda\mu\cdot}$ 等(6 个)，$A^{\lambda}_{\cdot\mu\nu\rho}$ 等(4 个)，$A_{\lambda\mu\nu\rho}$ 彼此亦为相伴张量. 其次应注意到，一般来说，

$$A^{\mu}_{\cdot\nu} \neq A_{\nu\cdot}^{\cdot\mu},\quad A_{\lambda\cdot\cdot\cdot}^{\cdot\mu\nu\rho} \neq A^{\mu\cdot\nu\rho}_{\cdot\lambda\cdot\cdot},$$

除非 $A^{\mu\nu}$，$A^{\lambda\mu\nu\rho}$ 对相应指标为对称张量.

由此可见，**指标排列次序的重要性**. 这里不仅指反变指标本身之间的次序和协变指标本身之间的次序是重要的，因为它直接涉及张量的对称性；而且还指反变指标与协变指标作为一个总体混排时，相互之间的次序所具有的重要性.

顺便指出，度规张量指标的升降定义混变度规张量为(由于 $g_{\mu\nu}$ 是对称张量)

$$g_{\mu\cdot}^{\cdot\nu} = g^{\nu}_{\cdot\mu} = \delta^{\nu}_{\cdot\mu} = \delta_{\mu\cdot}^{\cdot\nu}.$$

5.8.5　张量方程中的指标定则

为检验张量方程中指标的正确性，有以下一些简单定则.

(i) 整个方程所有各项中的自由指标必须匹配；即若它在一项中作为上标(下标)出现，则必须在所有各项中均作为上标(下标)出现.

(ii) 方程每项中的傀标必须单独匹配；即若它在某一项中出现时，必须作为上标和下标同时出现.

(iii) 任一项中任何指标都不应出现超过两次.

5.9 标量密度与张量密度

5.9.1 标量密度

在广义坐标变换下，凡具有下列变换性质者：

$$\varphi' = \left[\det\left\{\frac{\partial x'^{\alpha}}{\partial x^{\mu}}\right\}\right]^{w}\varphi,$$

称为权重为 w 的**标量密度**，它是 Descartes 张量中赝标量的推广.

例如，广义坐标变换下，P 点的体积元变换为

$$\begin{aligned}\mathrm{d}x^1\cdots\mathrm{d}x^n \to \mathrm{d}x'^1\cdots\mathrm{d}x'^n &= \det\left\{\frac{\partial x'^{\alpha}}{\partial x^{\mu}}\right\}\mathrm{d}x^1\cdots\mathrm{d}x^n \\ &= (\det\mathscr{A})\mathrm{d}x^1\cdots\mathrm{d}x^n,\end{aligned}$$

这是权重为 $+1$ 的标量密度的原型；这里曾令

$$\mathscr{A} = \{a^{\alpha}{}_{\mu}\},\quad a^{\alpha}{}_{\mu} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}}.$$

因为在 P 点的小体积 Δ 内的总质量 M 是不变量，所以，

$$M = \int_{\Delta}\rho\,\mathrm{d}x^1\cdots\mathrm{d}x^n = \int_{\Delta'}\rho'\mathrm{d}x'^1\cdots\mathrm{d}x'^n;$$

于是，

$$\rho' = (\det\mathscr{A})^{-1}\rho;$$

即，质量密度是权重为 -1 的标量密度.

同时，对于度规张量 $g_{\mu\nu}$ 的变换性质是

$$g'_{\alpha\beta} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\beta}}g_{\mu\nu};$$

引进矩阵记号

$$\mathscr{G} = \{g_{\mu\nu}\},\quad \mathscr{G}' = \{g'_{\alpha\beta}\};\quad \mathscr{B} = \{b^{\mu}{}_{\alpha}\},\quad b^{\mu}{}_{\alpha} \equiv \frac{\partial x^{\mu}}{\partial x'^{\alpha}};$$

上式可以写成

$$\mathscr{G}' = \widetilde{\mathscr{B}}\mathscr{G}\mathscr{B}.$$

于是，

$$\begin{aligned}\det\mathscr{G}' &= (\det\widetilde{\mathscr{B}})(\det\mathscr{G})(\det\mathscr{B}) \\ &= (\det\mathscr{A})^{-2}(\det\mathscr{G}),\end{aligned}$$

这里应用了 $\mathscr{B}=\mathscr{A}^{-1}$（见 5.6.4 节）以及 $(\det\mathscr{A}^{-1})=(\det\mathscr{A})^{-1}$（见第 3 章 3.4.2 节）；这是权重为 -2 的标量密度的原型.

由此可见（注意到 $\det\mathscr{G}=\det\{g_{\mu\nu}\}=g$），

$$\sqrt{|g'|}=\sqrt{|g|}(\det\mathscr{A})^{-1}.$$

5.9.2 标量积分元

将体积元的变换关系 $dx'^1\cdots dx'^n=(\det\mathscr{A})\,dx^1\cdots dx^n$ 与 $\sqrt{|g'|}=\sqrt{|g|}(\det\mathscr{A})^{-1}$相乘得到

$$\sqrt{|g'|}dx'^1\cdots dx'^n=\sqrt{|g|}dx^1\cdots dx^n.$$

它在广义坐标变换下为不变量,称为**标量积分元**.

5.9.3 张量密度

广义坐标变换下,凡其分量具有下列变换性质者:

$$\mathscr{T}'^{\alpha\beta\cdots}_{\cdots\cdots\xi\eta\cdots}=\left[\det\left\{\frac{\partial x'^\gamma}{\partial x^\lambda}\right\}\right]^w\frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x'^\beta}{\partial x^\nu}\cdots\frac{\partial x^\rho}{\partial x'^\xi}\frac{\partial x^\sigma}{\partial x'^\eta}\cdots\mathscr{T}^{\mu\nu\cdots}_{\cdots\cdots\rho\sigma\cdots},$$

称为权重为 w 的张量密度,它是 Descartes 张量中赝张量的推广.

5.10 商定律[①]

前面已经指出,两个张量的内积和外积,结果均为张量.然而,假定已知两个因子的乘积是张量,而且其中一个因子是张量.那么,能否断定另一因子也是张量呢?关于这个问题,有下述定律.

商定律 若一给定元素集合与一任意张量之(内或外)积已知为一张量,则此给定元素是一张量的分量.

用一个特例就足以证明此定律,因为下面的论证具有一般适用性.假定已知 $A(\lambda,\mu,\nu)$与任意张量 B^ρ_σ 之内积

$$A(\lambda,\mu,\nu)B^\lambda_\sigma=C^\nu_{\mu\sigma},$$

对一切这种 B^ρ_σ 均为张量.于是,变换至新坐标系后,上式变为

$$A'(\alpha,\beta,\gamma)B'^\alpha_\eta=C'^\gamma_{\beta\eta};$$

由于已知 B'^α_η 和 $C'^\gamma_{\beta\eta}$ 为张量,故有

$$\begin{aligned}A'(\alpha,\beta,\gamma)\frac{\partial x'^\alpha}{\partial x^\lambda}\frac{\partial x^\sigma}{\partial x'^\eta}B^\lambda_\sigma&=\frac{\partial x'^\gamma}{\partial x^\nu}\frac{\partial x^\mu}{\partial x'^\beta}\frac{\partial x^\sigma}{\partial x'^\eta}C^\nu_{\mu\sigma}\\&=\frac{\partial x'^\gamma}{\partial x^\nu}\frac{\partial x^\mu}{\partial x'^\beta}\frac{\partial x^\sigma}{\partial x'^\eta}A(\lambda,\mu,\nu)B^\lambda_\sigma,\end{aligned}$$

① 参考:D. F. Lawden, *An Introduction to Tensor Calculus, Relativity and Cosmology*, p. 94.

或者，

$$\left[A'(\alpha,\beta,\gamma)\frac{\partial x'^{\alpha}}{\partial x^{\lambda}}-\frac{\partial x'^{\gamma}}{\partial x^{\nu}}\frac{\partial x^{\mu}}{\partial x'^{\beta}}A(\lambda,\mu,\nu)\right]\frac{\partial x^{\sigma}}{\partial x'^{\eta}}B_{\sigma}^{\lambda}=0;$$

因为 B_{σ}^{λ} 为任意张量，总可选择使之当 $\lambda=\Lambda,\sigma=\Sigma$ 时有$\frac{\partial x^{\sigma}}{\partial x'^{\eta}}B_{\sigma}^{\lambda}=1$ 而所有其他情况下有 $B_{\sigma}^{\lambda}=0$，所以当 $\lambda=\Lambda$ 时前面方括号内表达式为零. 但因令 $\Lambda=1,2,\cdots,n$ 时此结果皆成立，故一般地有

$$A'(\alpha,\beta,\gamma)\frac{\partial x'^{\alpha}}{\partial x^{\lambda}}=\frac{\partial x'^{\gamma}}{\partial x^{\nu}}\frac{\partial x^{\mu}}{\partial x'^{\beta}}A(\lambda,\mu,\nu);\quad 或者$$

$$A'(\alpha,\beta,\gamma)=\frac{\partial x'^{\gamma}}{\partial x^{\nu}}\frac{\partial x^{\lambda}}{\partial x'^{\alpha}}\frac{\partial x^{\mu}}{\partial x'^{\beta}}A(\lambda,\mu,\nu),$$

这正是张量变换关系，表明

$$A(\lambda,\mu,\nu)=A_{\lambda\mu}^{\nu}$$

是反变一阶协变二阶的混变三阶张量.

下面考虑此定律的应用.

例如，令 A^{λ} 为任意反变矢量，则

$$\delta_{\mu}^{\lambda}A^{\mu}=A^{\lambda};$$

因为右边 A^{λ} 肯定为一矢量，因而根据商定律，δ_{μ}^{λ} 是混变二阶张量，这是 5.6.5 节中早已直接证明过的.

又例如，令 $g_{\mu\nu}$ 为对称协变张量，以及令 $g=|g_{\mu\nu}|\neq 0$ 为以张量分量为元素的行列式，并且以 $G^{\mu\nu}$ 表示此行列式中元素 $g_{\mu\nu}$ 的代数余子式. 于是，由第 3 章 3.3.2 节之推论 2 知，

$$g_{\mu\nu}G^{\lambda\nu}=g\delta_{\mu}^{\lambda},\quad g_{\mu\nu}G^{\mu\lambda}=g\delta_{\nu}^{\lambda}.$$

若令 A^{μ} 为任意反变矢量，并由

$$B_{\mu}=g_{\mu\nu}A^{\nu}$$

定义协变矢量 B_{μ}. 因为 $g\neq 0$，当任意选择 B_{μ} 的分量时，总可从最后这个方程计算出 A^{μ} 的相应分量，即 B_{μ} 和 A^{μ} 一起是任意的. 由于

$$\frac{1}{g}G^{\mu\nu}B_{\mu}=\frac{1}{g}G^{\mu\nu}g_{\mu\lambda}A^{\lambda}=\delta_{\lambda}^{\nu}A^{\lambda}=A^{\nu},$$

根据商定律，$\frac{1}{g}G^{\mu\nu}\equiv g^{\mu\nu}$ 是对称反变张量；因而

$$g_{\mu\nu}g^{\lambda\nu}=\delta_{\mu}^{\lambda},\quad g_{\mu\nu}g^{\mu\lambda}=\delta_{\nu}^{\lambda}.$$

相互之间具有上述关系的两个张量 $g^{\mu\nu}$ 和 $g_{\mu\nu}$ 称为**共轭张量**.

5.11　张量的微分运算[1][2]

为了进行张量的微分运算，首先引进下列简化记号，用逗号后带指标表示普通偏导数，并约定旧坐标系中指标用 $\lambda,\mu,\nu,\rho,\sigma$ 等表示，新坐标系中指标则用 $\alpha,\beta,\gamma,\xi,\eta$ 等表示. 例如，

$$x^{\mu}{}_{,\alpha} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}, \quad x'^{\alpha}{}_{,\mu} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}}, \quad x^{\mu}{}_{,\alpha\beta} = \frac{\partial^2 x^{\mu}}{\partial x'^{\alpha}\partial x'^{\beta}};$$

$$x^{\mu}{}_{,\alpha}x'^{\alpha}{}_{,\nu} = \delta^{\mu}_{\cdot\nu}, \quad x'^{\alpha}{}_{,\mu}x^{\mu}{}_{,\beta} = \delta^{\alpha}_{\cdot\beta};$$

$$\phi_{,\mu} = \frac{\partial\phi}{\partial x^{\mu}}, \quad \phi_{,\mu\nu} = \frac{\partial^2\phi}{\partial x^{\mu}\partial x^{\nu}}, \quad A^{\mu\nu}_{\cdot\cdot\lambda,\mu} = \partial A^{\mu\nu}_{\cdot\cdot\lambda}/\partial x^{\mu}.$$

必须指出，张量(标量除外)对坐标的普通偏导数并非张量. 例如，对 $A'_{\alpha} = x^{\mu}{}_{,\alpha}A_{\mu}$ 求导得

$$\begin{aligned} A'_{\alpha,\beta} &= x^{\mu}{}_{,\alpha\beta}A_{\mu} + x^{\mu}{}_{,\alpha}A_{\mu,\beta} \\ &= x^{\mu}{}_{,\alpha\beta}A_{\mu} + x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}A_{\mu,\nu}, \end{aligned}$$

由于多了右边第一项，故 $A_{\mu,\nu}$ 非张量，一般张量的导数亦如此.

5.11.1　矢量平移与仿射联络

矢量的普通偏导数不是张量，其原因很清楚：求导涉及两邻点矢量值之比较，而在广义坐标变换下，只有隶属于同一场点的矢量或张量才能进行代数运算，其结果方始为张量. 这就导致引进所谓矢量**平行位移**(简称**平移**)的概念.

令 P,Q 为空间中两邻近点 $x^{\mu}, x^{\mu}+\mathrm{d}x^{\mu}$，并令 $A_{\mu}, A_{\mu}+\mathrm{d}A_{\mu}$ 分别为与这两点相联系的协变矢量场之相应矢量. 假定将 P 点处矢量 A_{μ} 在其**长度或方向无改变**的情况下位移至邻近点 Q，则可视其为定义于 Q 点处的同一矢量，记为 $A_{\mu}+\delta A_{\mu}$. 于是，Q 点处两矢量 $A_{\mu}+\mathrm{d}A_{\mu}$ 与 $A_{\mu}+\delta A_{\mu}$ 之差 $\mathrm{d}A_{\mu}-\delta A_{\mu}$ 为矢量. 然而，在广义空间中，矢量的长度或方向无改变尚无明确含义.

若此空间为 Euclid 空间，并假定 y^k 为此空间中的直角坐标，而 B_k 为矢量场 A_{μ} 相对于这些直角坐标轴的分量. 现在，在此平直空间中，可将 P 点矢量 B_k 平行位移至 Q 点，其大小与方向均无改变，故可视为自由矢量，此时有 $B_k+\delta B_k = B_k$，即 $\delta B_k = 0$. 然而，后者在曲线坐标系 x^{μ} 下则应为 $A_{\mu}+\delta A_{\mu}$. 于是有

$$A_{\mu} = \frac{\partial y^k}{\partial x^{\mu}}B_k, \quad B_k = \frac{\partial x^{\lambda}}{\partial y^k}A_{\lambda},$$

---

① 参考：D. F. Lawden, *An Introduction to Tensor Calculus, Relativity and Cosmology*.

② 参考：A. W. Joshi, *Matrices and Tensors in Physics*.

$$\delta A_\mu = \delta\left(\frac{\partial y^k}{\partial x^\mu}B_k\right) = \delta\left(\frac{\partial y^k}{\partial x^\mu}\right)B_k = \frac{\partial^2 y^k}{\partial x^\mu \partial x^\nu}\mathrm{d}x^\nu \frac{\partial x^\lambda}{\partial y^k}A_\lambda;$$

结果是

$$\delta A_\mu = \Gamma^\lambda_{\mu\nu}A_\lambda \mathrm{d}x^\nu,$$

$$\Gamma^\lambda_{\mu\nu} = \frac{\partial^2 y^k}{\partial x^\mu \partial x^\nu}\frac{\partial x^\lambda}{\partial y^k}.$$

这表明，在 Euclid 空间中，δA_μ 是 A_λ 和 $\mathrm{d}x^\nu$ 的**双线性型**. 在广义空间中，可仿此由上述方程来定义 δA_μ，任意地确定空间每点处的 n^3 个量 $\Gamma^\lambda_{\mu\nu}$.（假定 $\Gamma^\lambda_{\mu\nu}$ 为 x^μ 的连续函数并具有所必需的各阶偏导数.）系数 $\Gamma^\lambda_{\mu\nu}$ 明确规定空间各点之间的一个**仿射联络**，又称**线性联络**，而 $\Gamma^\lambda_{\mu\nu}$ 则称为**仿射联络系数**.

仿射联络的一个最重要的特殊情形是 Riemann 空间的 Levi-Civita 联络，将在下面讨论.

5.11.2 Levi-Civita 联络①

(1) 矢量的长度

对于 Riemann 空间，有度规张量 $g_{\mu\nu}$，两邻近点 x^μ 和 $x^\mu+\mathrm{d}x^\mu$ 之间的距离平方表达为

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu.$$

位移 $\mathrm{d}x^\mu$ 是反变矢量的原型，故可仿此由

$$A^2 = g_{\mu\nu}A^\mu A^\nu = A_\nu A^\nu$$

来定义任意反变**矢量** A^μ **的长度** A，它是标量. 同时，还可由两矢量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的标积

$$\boldsymbol{A}\cdot\boldsymbol{B} = A_\mu B^\mu = A^\mu B_\mu$$

以及各自的长度 A 和 B 来定义**两矢量之间的夹角**：

$$\theta = \arccos\frac{A_\lambda B^\lambda}{\sqrt{(A_\mu A^\mu)(B_\nu B^\nu)}}.$$

这些定义与 Euclid 空间中的相应定义（参考 2.5.8 节）完全一致.

(2) Levi-Civita 平移

保持矢量长度不变的平行位移称为 **Levi-Civita 平移**.

容易证明，反变矢量 A^μ 由 P 点平移至 Q 点后变为 A^μ_Q，结果有[参考 5.11.3 节之(3)]：

$$A^\mu_Q = A^\mu + \delta A^\mu = A^\mu - \Gamma^\mu_{\lambda\nu}A^\lambda \mathrm{d}x^\nu,$$

右边均取 P 点处之值. 于是，Levi-Civita 平移可表达为：

$$g_{\mu\nu Q}A^\mu_Q A^\nu_Q = g_{\mu\nu}A^\mu A^\nu.$$

① 参考：俞允强，《广义相对论引论》，第二版，北京大学出版社，北京，1997.

将此表达式左边展开至 $\mathrm{d}x^\lambda$ 的一阶项为

$$(g_{\mu\nu}+g_{\mu\nu,\lambda}\mathrm{d}x^\lambda)(A^\mu-\Gamma^\mu_{\rho\sigma}A^\rho\mathrm{d}x^\sigma)(A^\nu-\Gamma^\nu_{\omega\tau}A^\omega\mathrm{d}x^\tau)$$

$$=g_{\mu\nu}A^\mu A^\nu+g_{\mu\nu,\lambda}A^\mu A^\nu\mathrm{d}x^\lambda-g_{\mu\nu}\Gamma^\mu_{\rho\sigma}A^\rho A^\nu\mathrm{d}x^\sigma-g_{\mu\nu}\Gamma^\nu_{\omega\tau}A^\omega A^\mu\mathrm{d}x^\tau.$$

在展开式右边第三项令傀标 $\sigma\to\lambda,\mu\to\sigma,\rho\to\mu$，第四项令 $\tau\to\lambda,\nu\to\sigma,\omega\to\nu$；将上式代入前式并消去第一项后立即得到

$$(g_{\mu\nu,\lambda}-g_{\sigma\nu}\Gamma^\sigma_{\mu\lambda}-g_{\mu\sigma}\Gamma^\sigma_{\nu\lambda})A^\mu A^\nu\mathrm{d}x^\lambda=0.$$

注意到 $A^\mu,\mathrm{d}x^\lambda$ 为任意矢量，于是有

$$g_{\mu\nu,\lambda}=g_{\sigma\nu}\Gamma^\sigma_{\mu\lambda}+g_{\mu\sigma}\Gamma^\sigma_{\nu\lambda}.$$

(3) Levi-Civita 联络

对上述方程作轮换 $(\lambda\mu\nu)$，$(\lambda\nu\mu)$ 得两新方程，连同原方程一起得到对于联络系数 $\Gamma^\sigma_{\mu\lambda}$ 等的线性方程组：

$$g_{\mu\nu,\lambda}=g_{\sigma\nu}\Gamma^\sigma_{\mu\lambda}+g_{\mu\sigma}\Gamma^\sigma_{\nu\lambda},$$
$$g_{\nu\lambda,\mu}=g_{\sigma\lambda}\Gamma^\sigma_{\nu\mu}+g_{\nu\sigma}\Gamma^\sigma_{\lambda\mu},$$
$$g_{\lambda\mu,\nu}=g_{\sigma\mu}\Gamma^\sigma_{\lambda\nu}+g_{\lambda\sigma}\Gamma^\sigma_{\mu\nu};$$

由于 $g_{\mu\nu}$ 是对称张量，上面共有 $\frac{1}{2}n^2(n+1)$ 个方程，而 $\Gamma^\lambda_{\mu\nu}$ 有 n^3 个独立分量，必须规定附加条件才能确定. 如果由后两式相加并减去第一式后再除以 2，记为 $[\mu\nu,\lambda]$，结果得到

$$[\mu\nu,\lambda]\equiv\frac{1}{2}(g_{\nu\lambda,\mu}+g_{\lambda\mu,\nu}-g_{\mu\nu,\lambda})$$
$$=g_{\lambda\sigma}\Gamma^\sigma_{(\mu\nu)}+g_{\mu\sigma}T^\sigma_{\lambda\nu}+g_{\nu\sigma}T^\sigma_{\lambda\mu};$$

这里 $\Gamma^\sigma_{(\mu\nu)}=\frac{1}{2}(\Gamma^\sigma_{\mu\nu}+\Gamma^\sigma_{\nu\mu})$ 是联络系数的对称组合，而 $T^\sigma_{\lambda\nu}=\frac{1}{2}(\Gamma^\sigma_{\lambda\nu}-\Gamma^\sigma_{\nu\lambda})$ 是其反对称组合；可以证明，后者是一张量，称为**挠率张量**(参见 5.11.6 节之(2)).

由此可见，如果限于讨论挠率为零的 Riemann 空间，联络系数是对称的，完全由度规场确定：

$$\Gamma_{\lambda\mu\nu}\equiv g_{\lambda\sigma}\Gamma^\sigma_{\mu\nu}=[\mu\nu,\lambda];$$

或者，两边缩乘以 $g^{\rho\lambda}$，并注意到 $g^{\rho\lambda}g_{\lambda\sigma}=\delta^\rho{}_\sigma$，最后得到对称的联络系数为

$$\Gamma^\rho_{\mu\nu}=g^{\rho\lambda}[\mu\nu,\lambda]\equiv\begin{Bmatrix}\rho\\ \mu\nu\end{Bmatrix};$$

$[\mu\nu,\lambda]$ 和 $\begin{Bmatrix}\rho\\ \mu\nu\end{Bmatrix}$ 分别称为**第一类和第二类 Christoffel[三指标]符号，简称 Christoffel 符号.**

由于 $g_{\mu\nu}$ 的对称性，显然有

$$[\mu\nu,\lambda]=[\nu\mu,\lambda],\begin{Bmatrix}\rho\\ \mu\nu\end{Bmatrix}=\begin{Bmatrix}\rho\\ \nu\mu\end{Bmatrix},g_{\mu\nu,\lambda}=[\mu\lambda,\nu]+[\nu\lambda,\mu].$$

Riemann 空间中满足① 挠率为零和② 矢量平移保持长度不变的仿射联络称为 **Levi-Civita 联络**，又称 **Riemann 联络**，有时亦称 **Christoffel 联络**.

5.11.3 张量的协变导数

(1) 协变张量的协变导数

根据前两小节的讨论，现在可以写出

$$dA_\mu - \delta A_\mu = A_{\mu,\nu}dx^\nu - \Gamma^\lambda_{\mu\nu}A_\lambda dx^\nu = (A_{\mu,\nu} - \Gamma^\lambda_{\mu\nu}A_\lambda)dx^\nu.$$

上面已指出，左边的 $dA_\mu - \delta A_\mu$ 是协变矢量，右边括号后因子为任意反变矢量，所以，根据商定律，括号内因子为协变二阶张量，称为 A_μ 的**协变导数**，用分号后带指标表示而记作 $A_{\mu;\nu}$，即

$$A_{\mu;\nu} = A_{\mu,\nu} - \Gamma^\lambda_{\mu\nu}A_\lambda.$$

下面将把求协变导数的过程推广至各阶各类型张量.

(2) 标量的梯度

首先考虑标量场 φ. 当经受平行位移时，其值不变，即总有 $\delta\varphi=0$. 因而

$$d\varphi - \delta\varphi = \varphi_{,\mu}dx^\mu,$$

结果有

$$\varphi_{;\mu} = \varphi_{,\mu},$$

即，**标量的协变导数与其普通偏导数或梯度相同**. 它是协变矢量，这是 5.6.4 节直接证明过的.

(3) 反变矢量的协变导数

现在令 B^μ 为反变矢量场，而 A_μ 为任意协变矢量，于是 $A_\mu B^\mu$ 是不变量. 在作平移时有

$$\delta(A_\mu B^\mu) = 0 \quad 或 \quad (\delta A_\mu)B^\mu + A_\mu \delta B^\mu = 0;$$

对于协变矢量有 $\delta A_\mu = \Gamma^\lambda_{\mu\nu}A_\lambda dx^\nu$，故

$$A_\lambda \delta B^\lambda = -\Gamma^\lambda_{\mu\nu}A_\lambda dx^\nu B^\mu.$$

因为 A_λ 是任意矢量，故有

$$\delta B^\lambda = -\Gamma^\lambda_{\mu\nu}B^\mu dx^\nu;$$

这个方程定义反变矢量的平移. 因此，反变矢量的协变导数为

$$B^\lambda_{;\nu} = B^\lambda_{,\nu} + \Gamma^\lambda_{\mu\nu}B^\mu.$$

(4) 混变二阶张量的协变导数

类似地，若 A^μ_ν 是张量场，B_μ，C^ν 是任意矢量，作平行位移时有

$$\delta(A^\mu_\nu B_\mu C^\nu) = 0;$$

容易推出

$$\delta A^\mu_\nu = \Gamma^\lambda_{\nu\sigma}A^\mu_\lambda dx^\sigma - \Gamma^\mu_{\lambda\sigma}A^\lambda_\nu dx^\sigma;$$

结果得到 A^{μ}_{ν} 的协变导数为

$$A^{\mu}_{\nu;\sigma} = A^{\mu}_{\nu,\sigma} - \Gamma^{\lambda}_{\nu\sigma}A^{\mu}_{\lambda} + \Gamma^{\mu}_{\lambda\sigma}A^{\lambda}_{\nu}.$$

(5) 求张量之协变导数的一般规则

容易看出，求任何张量之协变导数的规则可表述如下. 首先写下适当的普通偏导数，继之以各仿射联络项；后者通过下列方式获得：相对于张量的每个指标依次写下张量与仿射联络系数之内积，并且对应反变指标的项前加正号，对应协变指标的项前加负号. 或者，用公式写出为

$$A^{\lambda_1\cdots\lambda_l}_{\mu_1\cdots\mu_m;\nu} \equiv A^{\lambda_1\cdots\lambda_l}_{\mu_1\cdots\mu_m,\nu} + \sum_{r=1}^{l} A^{\lambda_1\cdots\lambda_{r-1}\rho\lambda_{r+1}\cdots\lambda_l}_{\mu_1\cdots\mu_m}\Gamma^{\lambda_r}_{\rho\nu} - \sum_{s=1}^{m} A^{\lambda_1\cdots\lambda_l}_{\mu_1\cdots\mu_{s-1}\sigma\mu_{s+1}\cdots\mu_m}\Gamma^{\sigma}_{\mu_s\nu}.$$

应该注意到，张量的协变导数为张量，其协变阶数增加一阶.

(6) 张量代数式的协变导数

由于张量的协变导数公式对该张量为线性的，不难得出张量和的协变导数为各张量协变导数之和. 例如，若 $C^{\mu}_{\nu}=A^{\mu}_{\nu}+B^{\mu}_{\nu}$，则有

$$C^{\mu}_{\nu;\lambda} = A^{\mu}_{\nu;\lambda} + B^{\mu}_{\nu;\lambda}.$$

另外，不难证明，对乘积的逐个因子求导数的公式，在协变导数中同样适用. 例如

$$\begin{aligned}(A^{\mu}_{\nu}B^{\nu})_{;\lambda} &= (A^{\mu}_{\nu}B^{\nu})_{,\lambda} + \Gamma^{\mu}_{\rho\lambda}(A^{\rho}_{\nu}B^{\nu}) \\ &= (A^{\mu}_{\nu,\lambda} + \Gamma^{\mu}_{\rho\lambda}A^{\rho}_{\nu} - \Gamma^{\rho}_{\nu\lambda}A^{\mu}_{\rho})B^{\nu} + (B^{\nu}_{,\lambda} + \Gamma^{\nu}_{\rho\lambda}B^{\rho})A^{\mu}_{\nu} \\ &= A^{\mu}_{\nu;\lambda}B^{\nu} + A^{\mu}_{\nu}B^{\nu}_{;\lambda}.\end{aligned}$$

(7) 两个特例

(i) **单位张量**(Kronecker δ 符号)$\delta^{\mu}{}_{\nu}$ **的协变导数为零**. 因为

$$\delta^{\mu}{}_{\nu;\lambda} = \delta^{\mu}{}_{\nu,\lambda} + \delta^{\rho}{}_{\nu}\Gamma^{\mu}_{\rho\lambda} - \delta^{\mu}{}_{\rho}\Gamma^{\rho}_{\nu\lambda} = \Gamma^{\mu}_{\nu\lambda} - \Gamma^{\mu}_{\nu\lambda} = 0;$$

这里应用了其普通偏导数 $\delta^{\mu}{}_{\nu,\lambda}=0$.

(ii) **度规张量** $g_{\mu\nu}$ **或** $g^{\mu\nu}$ **的协变导数为零.**

对于协变度规张量，其协变导数

$$g_{\mu\nu;\lambda} = g_{\mu\nu,\lambda} - g_{\sigma\nu}\Gamma^{\sigma}_{\mu\lambda} - g_{\mu\sigma}\Gamma^{\sigma}_{\nu\lambda} = 0,$$

右边为零是 5.11.2 节之(2)中最后一式的结果.

对于反变度规张量 $g^{\mu\nu}$ 的协变导数，可从

$$\delta^{\nu}{}_{\sigma} = g^{\rho\nu}g_{\sigma\rho}$$

出发，利用对乘积求导公式，对其求协变导数并乘以 $g^{\sigma\mu}$ 得

$$g^{\sigma\mu}\delta^{\nu}{}_{\sigma;\lambda} = g^{\sigma\mu}(g_{\sigma\rho}g^{\rho\nu}{}_{;\lambda} + g^{\rho\nu}g_{\sigma\rho;\lambda}) = g^{\mu\nu}{}_{;\lambda} + g^{\sigma\mu}g^{\rho\nu}g_{\sigma\rho;\lambda};$$

由于前面已得 $\delta^{\nu}{}_{\sigma;\lambda}=0$，$g_{\sigma\rho;\lambda}=0$，故有

$$g^{\mu\nu}{}_{;\lambda} = 0.$$

顺便提一句，对于 $g^{\mu\nu}$ 的普通偏导数，显然有

$$g^{\mu\nu}{}_{,\lambda} = - g^{\sigma\mu} g^{\rho\nu} g_{\sigma\rho,\lambda}.$$

由于 $g_{\mu\nu;\lambda}=0$ 和 $g^{\mu\nu}{}_{;\lambda}=0$，因此，可将 $g_{\mu\nu}$ 或 $g^{\mu\nu}$ 从表示协变导数的分号前的括号中任意拉出或拉进.

5.11.4 张量的协变散度

将张量的协变导数中分号后指标与任一反变指标进行缩并，称为张量的**协变散度**.

例如，对于张量 $A^{\mu\nu}{}_{\lambda}$ 可构成两个协变散度

$$\mathrm{div}_{\mu} A^{\mu\nu}{}_{\lambda} = A^{\mu\nu}{}_{\lambda;\mu} \quad 或 \quad \mathrm{div}_{\nu} A^{\mu\nu}{}_{\lambda} = A^{\mu\nu}{}_{\lambda;\nu}.$$

下面采用 Levi-Civita 联络. 因为要用到 Christoffel 符号的缩并，首先推导其具体表达式.

(1) Christoffel 符号的缩并

因为行列式的导数可通过分别对其每行求导再对此结果求和而获得[参见第 3 章 3.9.1 节]；故有

$$g_{,\lambda} = G^{\mu\nu} g_{\mu\nu,\lambda} = g g^{\mu\nu} g_{\mu\nu,\lambda}.$$

因为由 5.11.2 节之(3)中有

$$g_{\mu\nu,\lambda} = [\mu\lambda,\nu] + [\nu\lambda,\mu],$$

$$g_{,\lambda} = g g^{\mu\nu}([\mu\lambda,\nu] + [\nu\lambda,\mu]) = 2g\Gamma^{\mu}_{\lambda\mu}, 或$$

$$\Gamma^{\mu}_{\lambda\mu} = \frac{1}{2} g^{\mu\nu} g_{\mu\nu,\lambda} = (\ln\sqrt{-g})_{,\lambda};$$

这里用 $-g$，因为物理中所用度规张量，其行列式常为负.

下面举出协变散度的一些具体例子.

(2) 反变矢量的协变散度

反变矢量 A^{μ} 的协变散度是标量，可写成

$$A^{\mu}{}_{;\mu} = A^{\mu}{}_{,\mu} + \Gamma^{\mu}_{\lambda\mu} A^{\lambda} = A^{\mu}{}_{,\mu} + (\ln\sqrt{-g})_{,\lambda} A^{\lambda} = \frac{1}{\sqrt{-g}} (A^{\mu}\sqrt{-g})_{,\mu}.$$

(3) Laplace 符的运算

若矢量场为标量场 φ 的梯度，即

$$A_{\mu} = \varphi_{,\mu} \quad 和 \quad A^{\mu} = g^{\mu\nu}\varphi_{,\nu},$$

则得此矢量的散度为

$$\mathrm{div\ grad}\ \varphi = \nabla^2\varphi = \frac{1}{\sqrt{-g}} (\sqrt{-g}\, g^{\mu\nu}\varphi_{,\nu})_{,\mu}.$$

此式右边代表 Riemann 空间中 **Laplace 符**对 φ 的**运算**.

(4) 反变反对称张量的协变散度

$$F^{\mu\nu}{}_{;\nu}=F^{\mu\nu}{}_{,\nu}+\Gamma^{\mu}_{\sigma\nu}F^{\sigma\nu}+\Gamma^{\nu}_{\sigma\nu}F^{\mu\sigma}=\frac{1}{\sqrt{-g}}(F^{\mu\nu}\sqrt{-g})_{,\nu};$$

因为第一行第二项傀标 $\sigma\leftrightarrow\nu$ 互换时，$\Gamma^{\mu}_{\nu\sigma}=\Gamma^{\mu}_{\sigma\nu}$ 而 $F^{\nu\sigma}=-F^{\sigma\nu}$，故有 $\Gamma^{\mu}_{\sigma\nu}F^{\sigma\nu}=\Gamma^{\mu}_{\nu\sigma}F^{\nu\sigma}=-\Gamma^{\mu}_{\sigma\nu}F^{\sigma\nu}=0$.

(5) 混变二阶张量的协变散度

$$N^{\nu}_{\mu;\nu}=N^{\nu}_{\mu,\nu}+\Gamma^{\nu}_{\sigma\nu}N^{\sigma}_{\mu}-\Gamma^{\sigma}_{\mu\nu}N^{\nu}_{\sigma}=\frac{1}{\sqrt{-g}}(N^{\nu}_{\mu}\sqrt{-g})_{,\nu}-\Gamma^{\sigma}_{\mu\nu}N^{\nu}_{\sigma};$$

上式常用于 $g^{\lambda\sigma}N^{\nu}_{\sigma}=N^{\lambda\nu}$ 为对称张量情况. 因为这时有

$$\begin{aligned}\Gamma^{\sigma}_{\mu\nu}N^{\nu}_{\sigma}&=g^{\sigma\lambda}[\mu\nu,\lambda]N^{\nu}_{\sigma}=N^{\lambda\nu}[\mu\nu,\lambda]=N^{\nu\lambda}[\mu\lambda,\nu]\\&=\frac{1}{2}N^{\lambda\nu}\{[\mu\nu,\lambda]+[\mu\lambda,\nu]\}=\frac{1}{2}N^{\lambda\nu}g_{\lambda\nu,\mu};\end{aligned}$$

所以，在此情况下，上式化为

$$N^{\nu}_{\mu;\nu}=\frac{1}{\sqrt{-g}}(N^{\nu}_{\mu}\sqrt{-g})_{,\nu}-\frac{1}{2}N^{\lambda\nu}g_{\lambda\nu,\mu}.$$

5.11.5 联络系数的变换律

(1) 一般性推导

协变矢量 A_{μ} 的协变导数 $A_{\mu;\nu}$ 是张量，它由

$$A_{\mu;\nu}=A_{\mu,\nu}-\Gamma^{\lambda}_{\mu\nu}A_{\lambda}$$

定义；此张量方程在广义坐标变换下具有不变性，即有

$$A'_{\alpha;\beta}=A'_{\alpha,\beta}-\Gamma'^{\gamma}_{\alpha\beta}A'_{\gamma};$$

现在可以通过张量的变换性质来推导联络系数 $\Gamma^{\lambda}_{\mu\nu}$ 所应遵循的变换律，我们有

$$\begin{aligned}x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}A_{\mu;\nu}&=(x^{\lambda}{}_{,\alpha}A_{\lambda})_{,\beta}-\Gamma'^{\gamma}_{\alpha\beta}x^{\lambda}{}_{,\gamma}A_{\lambda}\\&=x^{\lambda}{}_{,\alpha}x^{\nu}{}_{,\beta}A_{\lambda,\nu}+x^{\lambda}{}_{,\alpha\beta}A_{\lambda}-\Gamma'^{\gamma}_{\alpha\beta}x^{\lambda}{}_{,\gamma}A_{\lambda}.\end{aligned}$$

将 $A_{\mu;\nu}$ 的表达式代入，消去右边第一项后得到

$$-x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}\Gamma^{\lambda}_{\mu\nu}A_{\lambda}=x^{\lambda}{}_{,\alpha\beta}A_{\lambda}-\Gamma'^{\gamma}_{\alpha\beta}x^{\lambda}{}_{,\gamma}A_{\lambda}.$$

因为 A_{λ} 为任意矢量，所以有

$$\Gamma'^{\gamma}_{\alpha\beta}x^{\lambda}{}_{,\gamma}=x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}\Gamma^{\lambda}_{\mu\nu}+x^{\lambda}{}_{,\alpha\beta}.$$

两边同乘以 $x'^{\xi}{}_{,\lambda}$，注意到 $x'^{\xi}{}_{,\lambda}x^{\lambda}{}_{,\gamma}=\delta^{\xi}{}_{\gamma}$，结果有(再令 $\xi\rightarrow\gamma$)

$$\Gamma'^{\gamma}_{\alpha\beta}=x'^{\gamma}{}_{,\lambda}x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}\Gamma^{\lambda}_{\mu\nu}+x'^{\gamma}{}_{,\lambda}x^{\lambda}{}_{,\alpha\beta};$$

这就是**联络系数 $\Gamma^{\lambda}_{\mu\nu}$ 的变换律**. 可以看出，由于多了右边第二项，所以，$\Gamma^{\lambda}_{\mu\nu}$ 并非张量；但其变分为张量，因为

$$\delta\Gamma'^{\gamma}_{\alpha\beta}=x'^{\gamma}{}_{,\lambda}x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}\delta\Gamma^{\lambda}_{\mu\nu}.$$

(2) 由度规张量推导

在 Riemann 空间中，对度规张量的变换关系 $g'_{\alpha\beta}=x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}g_{\mu\nu}$ 求导，并作轮换 $(\alpha\beta\gamma)$，容易求得

$$g'_{\alpha\beta,\gamma}=[x^{\mu}{}_{,\alpha\gamma}x^{\nu}{}_{,\beta}+x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta\gamma}]g_{\mu\nu}+x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}x^{\sigma}{}_{,\gamma}g_{\mu\nu,\sigma},$$
$$g'_{\beta\gamma,\alpha}=[x^{\mu}{}_{,\beta\alpha}x^{\nu}{}_{,\gamma}+x^{\mu}{}_{,\beta}x^{\nu}{}_{,\gamma\alpha}]g_{\mu\nu}+x^{\mu}{}_{,\beta}x^{\nu}{}_{,\gamma}x^{\sigma}{}_{,\alpha}g_{\mu\nu,\sigma},$$
$$g'_{\gamma\alpha,\beta}=[x^{\mu}{}_{,\gamma\beta}x^{\nu}{}_{,\alpha}+x^{\mu}{}_{,\gamma}x^{\nu}{}_{,\alpha\beta}]g_{\mu\nu}+x^{\mu}{}_{,\gamma}x^{\nu}{}_{,\alpha}x^{\sigma}{}_{,\beta}g_{\mu\nu,\sigma};$$

在第一式右边第一项对傀标作对换$(\mu\nu)$，后两式右边第二项对傀标分别作替换 $\mu\to\nu\to\sigma\to\mu$ 和 $\mu\to\sigma\to\nu\to\mu$；最后，将后两式相加减去第一式后除以 2，并注意到 $g_{\mu\nu}$ 为对称张量和 $x^{\mu}{}_{,\beta\alpha}$ 等中逗号后指标次序不重要，则根据定义得第一类 Christoffel 符号的变换律：

$$[\alpha\beta,\gamma]'=x^{\mu}{}_{,\alpha\beta}x^{\nu}{}_{,\gamma}g_{\mu\nu}+x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}x^{\sigma}{}_{,\gamma}[\mu\nu,\sigma].$$

由于多了右边第一项，故$[\mu\nu,\sigma]$并非张量.

现在用 $g'^{\gamma\xi}$ 与上式进行缩乘，左边按定义变为 $\Gamma'^{\xi}_{\alpha\beta}$；右边第一项变为

$$x^{\mu}{}_{,\alpha\beta}x^{\nu}{}_{,\gamma}x'^{\gamma}{}_{,\rho}x'^{\xi}{}_{,\lambda}g^{\rho\lambda}g_{\mu\nu}=x^{\mu}{}_{,\alpha\beta}x'^{\xi}{}_{,\mu},$$

这里先用了 $x^{\nu}{}_{,\gamma}x'^{\gamma}{}_{,\rho}=\delta^{\nu}{}_{\cdot\rho}$，再用 $g^{\nu\lambda}g_{\mu\nu}=\delta_{\mu}{}^{\cdot\lambda}$；右边第二项变为

$$x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}x^{\sigma}{}_{,\gamma}x'^{\gamma}{}_{,\rho}x'^{\xi}{}_{,\lambda}g^{\rho\lambda}[\mu\nu,\sigma]=x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}x'^{\xi}{}_{,\lambda}\Gamma^{\lambda}_{\mu\nu},$$

这里用了 $x^{\sigma}{}_{,\gamma}x'^{\gamma}{}_{,\rho}=\delta^{\sigma}{}_{\cdot\rho}$ 和 $\Gamma^{\lambda}_{\mu\nu}$ 的定义；于是，最后(再令 $\xi\to\gamma$)得第二类 Christoffel 符号的变换律：

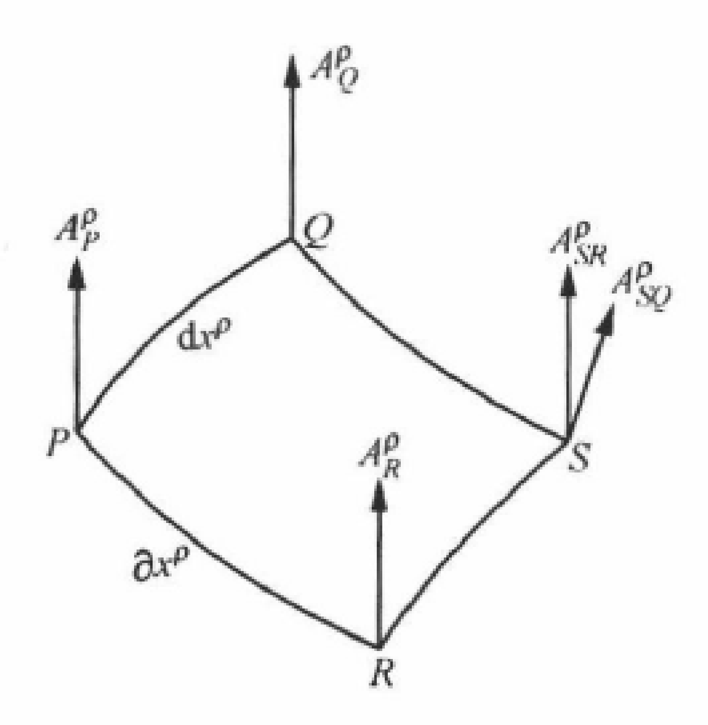
$$\Gamma'^{\gamma}_{\alpha\beta}=x^{\mu}{}_{,\alpha\beta}x'^{\gamma}{}_{,\mu}+x^{\mu}{}_{,\alpha}x^{\nu}{}_{,\beta}x'^{\gamma}{}_{,\lambda}\Gamma^{\lambda}_{\mu\nu};$$

这正是(1)中早已推导出的结果.

5.11.6 曲率张量

(1) 由矢量平移定义求

考虑如图所示无穷小平行四边形 $PQSR$，具有邻边 $PQ=\mathrm{d}x^{\rho}$，$PR=\partial x^{\rho}$. 定义于 P 点的反变矢量 A^{ρ}_{P} 经分别平移至 Q，R 再至 S 后的矢量相应为 A^{ρ}_{Q}，A^{ρ}_{R} 以及 A^{ρ}_{SQ} 和 A^{ρ}_{SR} 如下图所示.



根据平行位移定义可求得

$$\begin{aligned}A^{\rho}_{SQ} &= A^{\rho}_{Q} + \delta A^{\rho}_{Q} = A^{\rho}_{Q} - \Gamma^{\rho}_{\sigma\nu Q} A^{\sigma}_{Q} \partial x^{\nu} \\ &= (A^{\rho}_{P} - \Gamma^{\rho}_{\nu\lambda P} A^{\nu}_{P} \mathrm{d}x^{\lambda}) - (\Gamma^{\rho}_{\sigma\nu P} + \Gamma^{\rho}_{\sigma\nu P,\tau} \mathrm{d}x^{\tau})(A^{\sigma}_{P} - \Gamma^{\sigma}_{\lambda\mu P} A^{\lambda}_{P} \mathrm{d}x^{\mu}) \partial x^{\nu} \\ &= A^{\rho} - \Gamma^{\rho}_{\nu\lambda} A^{\nu} \mathrm{d}x^{\lambda} - \Gamma^{\rho}_{\sigma\nu} A^{\sigma} \partial x^{\nu} \\ &\quad - \Gamma^{\rho}_{\sigma\nu,\tau} A^{\sigma} \mathrm{d}x^{\tau} \partial x^{\nu} + \Gamma^{\rho}_{\sigma\nu} \Gamma^{\sigma}_{\lambda\mu} A^{\lambda} \mathrm{d}x^{\mu} \partial x^{\nu} + o[(\mathrm{d}x^{\lambda})^{3}];\end{aligned}$$

这里还用了 $\Gamma^{\rho}_{\sigma\nu Q}$ 在 P 点的展开，并在最后表达式中省略了下标 P.

类似地，可求得

$$\begin{aligned}A^{\rho}_{SR} &= A^{\rho} - \Gamma^{\rho}_{\nu\lambda} A^{\nu} \partial x^{\lambda} - \Gamma^{\rho}_{\sigma\nu} A^{\sigma} \mathrm{d}x^{\nu} \\ &\quad - \Gamma^{\rho}_{\sigma\nu,\tau} A^{\sigma} \partial x^{\tau} \mathrm{d}x^{\nu} + \Gamma^{\rho}_{\sigma\nu} \Gamma^{\sigma}_{\lambda\mu} A^{\lambda} \partial x^{\mu} \mathrm{d}x^{\nu} + o[(\mathrm{d}x^{\lambda})^{3}].\end{aligned}$$

以上两式相减，注意到傀标可换字母，两式右边前三项相消，至 $\mathrm{d}x^{\lambda}$ 的二次项给出

$$\begin{aligned}A^{\rho}_{SR} - A^{\rho}_{SQ} &= \Gamma^{\rho}_{\sigma\nu,\tau} A^{\sigma} \mathrm{d}x^{\tau} \partial x^{\nu} - \Gamma^{\rho}_{\sigma\nu,\tau} A^{\sigma} \partial x^{\tau} \mathrm{d}x^{\nu} \\ &\quad + \Gamma^{\rho}_{\sigma\nu} \Gamma^{\sigma}_{\lambda\mu} A^{\lambda} \partial x^{\mu} \mathrm{d}x^{\nu} - \Gamma^{\rho}_{\sigma\nu} \Gamma^{\sigma}_{\lambda\mu} A^{\lambda} \mathrm{d}x^{\mu} \partial x^{\nu}.\end{aligned}$$

在右边前两项作指标代换 $\sigma \to \lambda$ 和 $\tau \to \mu$，然后再在第二和第三项作交换 $\mu \leftrightarrow \nu$，于是，上述方程最后可写成

$$\begin{aligned}A^{\rho}_{SR} - A^{\rho}_{SQ} &= R^{\rho}_{\lambda\mu\nu} A^{\lambda} \mathrm{d}x^{\mu} \partial x^{\nu}, \\ R^{\rho}_{\lambda\mu\nu} &= \Gamma^{\rho}_{\lambda\nu,\mu} - \Gamma^{\rho}_{\lambda\mu,\nu} + \Gamma^{\rho}_{\sigma\mu} \Gamma^{\sigma}_{\lambda\nu} - \Gamma^{\rho}_{\sigma\nu} \Gamma^{\sigma}_{\lambda\mu};\end{aligned}$$

因为 $A^{\rho}_{SR} - A^{\rho}_{SQ}$ 是矢量，A^{λ}，$\mathrm{d}x^{\mu}$，∂x^{ν} 是任意矢量，根据商定律，$R^{\rho}_{\lambda\mu\nu}$ 是混变四阶张量，称为 Riemann-Christoffel **曲率张量**.

(2) 由二阶协变导数的交换律引进

还可通过求协变矢量场 A_{λ} 的二阶协变导数来引进曲率张量

$$\begin{aligned}A_{\lambda;\mu\nu} &\equiv (A_{\lambda;\mu})_{;\nu} = (A_{\lambda;\mu})_{,\nu} - \Gamma^{\sigma}_{\lambda\nu} A_{\sigma;\mu} - \Gamma^{\sigma}_{\mu\nu} A_{\lambda;\sigma} \\ &= A_{\lambda,\mu\nu} - \Gamma^{\rho}_{\lambda\mu,\nu} A_{\rho} - \Gamma^{\rho}_{\lambda\mu} A_{\rho,\nu} - \Gamma^{\sigma}_{\lambda\nu}(A_{\sigma,\mu} - \Gamma^{\rho}_{\sigma\mu} A_{\rho}) - \Gamma^{\sigma}_{\mu\nu}(A_{\lambda,\sigma} - \Gamma^{\rho}_{\lambda\sigma} A_{\rho}) \\ &= A_{\lambda,\mu\nu} - \Gamma^{\rho}_{\lambda\mu} A_{\rho,\nu} - \Gamma^{\sigma}_{\lambda\nu} A_{\sigma,\mu} - \Gamma^{\sigma}_{\mu\nu} A_{\lambda,\sigma} - \Gamma^{\rho}_{\lambda\mu,\nu} A_{\rho} + \Gamma^{\sigma}_{\lambda\nu} \Gamma^{\rho}_{\sigma\mu} A_{\rho} + \Gamma^{\sigma}_{\mu\nu} \Gamma^{\rho}_{\lambda\sigma} A_{\rho}, \\ A_{\lambda;\nu\mu} &= A_{\lambda,\nu\mu} - \Gamma^{\rho}_{\lambda\nu} A_{\rho,\mu} - \Gamma^{\sigma}_{\lambda\mu} A_{\sigma,\nu} - \Gamma^{\sigma}_{\nu\mu} A_{\lambda,\sigma} - \Gamma^{\rho}_{\lambda\nu,\mu} A_{\rho} + \Gamma^{\sigma}_{\lambda\mu} \Gamma^{\rho}_{\sigma\nu} A_{\rho} + \Gamma^{\sigma}_{\nu\mu} \Gamma^{\rho}_{\lambda\sigma} A_{\rho};\end{aligned}$$

两者相减，并注意到对于普通偏导数而言，求导次序可交换，即

$$A_{\lambda,\mu\nu} - A_{\lambda,\nu\mu} = 0,$$

以及傀标可换字母，结果得到协变导数(对求导次序)的交换律：

$$A_{\lambda;\mu\nu} - A_{\lambda;\nu\mu} = R^{\rho}_{\lambda\mu\nu} A_{\rho} - 2T^{\rho}_{\mu\nu} A_{\lambda;\rho};$$

其中

$$R^{\rho}_{\lambda\mu\nu} = \Gamma^{\rho}_{\lambda\nu,\mu} - \Gamma^{\rho}_{\lambda\mu,\nu} + \Gamma^{\sigma}_{\lambda\nu} \Gamma^{\rho}_{\sigma\mu} - \Gamma^{\sigma}_{\lambda\mu} \Gamma^{\rho}_{\sigma\nu}$$

正是前面求出过的**曲率张量**，而

$$T^{\rho}_{\mu\nu} = \frac{1}{2}(\Gamma^{\rho}_{\mu\nu} - \Gamma^{\rho}_{\nu\mu})$$

称为仿射空间的**挠率张量**.

下面将进一步讨论无挠率 Riemann 空间中的曲率张量.

(3) 协变曲率张量

用度规张量 $g_{\rho\tau}$ 对曲率张量

$$R^{\rho}_{\lambda\mu\nu} = \Gamma^{\rho}_{\lambda\nu,\mu} - \Gamma^{\rho}_{\lambda\mu,\nu} + \Gamma^{\sigma}_{\lambda\nu}\Gamma^{\rho}_{\sigma\mu} - \Gamma^{\sigma}_{\lambda\mu}\Gamma^{\rho}_{\sigma\nu}$$

进行缩乘得到

$$R_{\tau\lambda\mu\nu} = g_{\rho\tau}R^{\rho}_{\lambda\mu\nu},$$

称为**协变曲率张量**.

为了下面讨论曲率张量的对称性,这里先将协变曲率张量用度规张量的一阶和二阶导数,主要通过第一类 Christoffel 符号

$$[\mu\nu,\lambda] = \frac{1}{2}(g_{\lambda\nu,\mu} + g_{\lambda\mu,\nu} - g_{\mu\nu,\lambda})$$

及其导数来表达,还有 $g^{\mu\nu}$ 出现在系数中.

首先注意到

$$[\lambda\mu,\tau]_{,\nu} = (g_{\tau\rho}\Gamma^{\rho}_{\lambda\mu})_{,\nu} = g_{\tau\rho,\nu}\Gamma^{\rho}_{\lambda\mu} + g_{\tau\rho}\Gamma^{\rho}_{\lambda\mu,\nu},$$

于是,$R_{\tau\lambda\mu\nu}$ 可表达为

$$\begin{aligned}
R_{\tau\lambda\mu\nu} &= [\lambda\nu,\tau]_{,\mu} - g_{\tau\rho,\mu}\Gamma^{\rho}_{\lambda\nu} - [\lambda\mu,\tau]_{,\nu} + g_{\tau\rho,\nu}\Gamma^{\rho}_{\lambda\mu} + g_{\tau\rho}(\Gamma^{\sigma}_{\lambda\nu}\Gamma^{\rho}_{\sigma\mu} - \Gamma^{\sigma}_{\lambda\mu}\Gamma^{\rho}_{\sigma\nu}) \\
&= [\lambda\nu,\tau]_{,\mu} - [\lambda\mu,\tau]_{,\nu} + \Gamma^{\sigma}_{\lambda\mu}(g_{\tau\sigma,\nu} - [\sigma\nu,\tau]) - \Gamma^{\sigma}_{\lambda\nu}(g_{\tau\sigma,\mu} - [\sigma\mu,\tau]) \\
&= [\lambda\nu,\tau]_{,\mu} - [\lambda\mu,\tau]_{,\nu} + \Gamma^{\sigma}_{\lambda\mu}[\tau\nu,\sigma] - \Gamma^{\sigma}_{\lambda\nu}[\tau\mu,\sigma] \\
&= [\lambda\nu,\tau]_{,\mu} - [\lambda\mu,\tau]_{,\nu} + g^{\sigma\rho}\{[\lambda\mu,\rho][\tau\nu,\sigma] - [\lambda\nu,\rho][\tau\mu,\sigma]\}.
\end{aligned}$$

由此可见,$R_{\tau\lambda\mu\nu}$ 中各项,或含 $g_{\mu\nu}$ 的二阶导数一次,或含其一阶导数二次,还有 $g^{\sigma\rho}$ 出现在系数中.

(4) 曲率张量的对称性

由于度规张量 $g^{\mu\nu}$ 和联络系数 $\Gamma^{\lambda}_{\mu\nu}$ 的对称性,根据曲率张量的定义式,曲率张量具有下列四种类型的对称性.

(i) **曲率张量对最后两个协变指标的交换为反对称的**,即

$$R^{\rho}_{\lambda\mu\nu} = -R^{\rho}_{\lambda\nu\mu}; \quad R_{\tau\lambda\mu\nu} = -R_{\tau\lambda\nu\mu}.$$

关于这点,由其定义式立即可以看出.

(ii) **协变曲率张量对前两个指标的交换为反对称的**,即

$$R_{\tau\lambda\mu\nu} = -R_{\lambda\tau\mu\nu}.$$

这是因为,将(3)中 $R_{\tau\lambda\mu\nu}$ 的表达式之前两项展开后容易得到(注意到 $g_{\tau\nu,\mu\lambda}$ 中逗号前后两指标各自可交换)

$$[\lambda\nu,\tau]_{,\mu} - [\lambda\mu,\tau]_{,\nu} = \frac{1}{2}(g_{\tau\nu,\mu\lambda} - g_{\lambda\nu,\mu\tau} - g_{\tau\mu,\nu\lambda} + g_{\lambda\mu,\nu\tau})$$

$$=-\{[\tau\nu,\lambda]_{,\mu}-[\tau\mu,\lambda]_{,\nu}\},$$

它对 $\tau\leftrightarrow\lambda$ 为反对称的；另外，对表达式之后两项令傀标 $\sigma\leftrightarrow\rho$ 后容易得到

$$\begin{aligned}&g^{\sigma\rho}\{[\lambda\mu,\rho][\tau\nu,\sigma]-[\lambda\nu,\rho][\tau\mu,\sigma]\}\\&\quad=g^{\rho\sigma}\{[\lambda\mu,\sigma][\tau\nu,\rho]-[\lambda\nu,\sigma][\tau\mu,\rho]\}\\&\quad=-g^{\sigma\rho}\{[\tau\mu,\rho][\lambda\nu,\sigma]-[\tau\nu,\rho][\lambda\mu,\sigma]\},\end{aligned}$$

它对 $\tau\leftrightarrow\lambda$ 也为反对称的；所以，$R_{\tau\lambda\mu\nu}$ 对 $\tau\leftrightarrow\lambda$ 为反对称的.

（iii）**协变曲率张量对前后两对指标的交换为对称的**，即

$$R_{\tau\lambda\mu\nu}=R_{\mu\nu\tau\lambda}.$$

这是因为，在 $\tau\lambda\leftrightarrow\mu\nu$（即 $\tau\leftrightarrow\mu,\lambda\leftrightarrow\nu$）时，容易看出：

$$[\lambda\nu,\tau]_{,\mu}-[\lambda\mu,\tau]_{,\nu}=\frac{1}{2}(g_{\tau\nu,\mu\lambda}-g_{\lambda\nu,\mu\tau}-g_{\tau\mu,\nu\lambda}+g_{\lambda\mu,\nu\tau})$$

右边两正项交换，两负项不变；而

$$g^{\sigma\rho}\{[\lambda\mu,\rho][\tau\nu,\sigma]-[\lambda\nu,\rho][\tau\mu,\sigma]\}$$

中负项不变，正项再作傀标 $\sigma\leftrightarrow\rho$ 后亦不变；故具有上述对称性.

（iv）**协变曲率张量，若令其任一指标固定不变，而令其余指标作轮换，这些分量相加的结果为零**；例如，若第一个指标固定则有

$$R_{\tau\lambda\mu\nu}+R_{\tau\mu\nu\lambda}+R_{\tau\nu\lambda\mu}=0.$$

这是因为，$R_{\tau\lambda\mu\nu}$ 可表达为

$$\begin{aligned}R_{\tau\lambda\mu\nu}=&\frac{1}{2}(g_{\tau\nu,\mu\lambda}-g_{\lambda\nu,\mu\tau}-g_{\tau\mu,\nu\lambda}+g_{\lambda\mu,\nu\tau})\\&+g^{\sigma\rho}\{[\lambda\mu,\rho][\tau\nu,\sigma]-[\lambda\nu,\rho][\tau\mu,\sigma]\},\end{aligned}$$

将此式与轮换$(\lambda\mu\nu)$、$(\lambda\nu\mu)$后所得两式相加，立即可得其结果为零. 应该指出，分别令其他指标固定所得三个类似关系可利用前三个对称性变换得到；它们并非独立关系.

协变曲率张量分量间的这四个独立关系，可合并写成

$$\begin{aligned}R_{\tau\lambda\mu\nu}&=-R_{\lambda\tau\mu\nu}=-R_{\tau\lambda\nu\mu}=R_{\mu\nu\tau\lambda},\\R_{\tau\lambda\mu\nu}&+R_{\tau\mu\nu\lambda}+R_{\tau\nu\lambda\mu}=0.\end{aligned}$$

（5）协变曲率张量的独立分量数

由于前述两个反对称关系，曲率张量之形如 $R_{\lambda\lambda\mu\nu}$，$R_{\tau\lambda\mu\mu}$，$R_{\lambda\lambda\mu\mu}$ 和 $R_{\tau\tau\tau\tau}$（本段不采用求和约定）的分量均恒等于零. 另外，分量之间还需满足上述对称性关系. 这样使

得总共 n^4 个分量中所剩不为零的独立分量可分为下列三种类型①：

(i) 2 个不同指标，$R_{\tau\lambda\tau\lambda}\ (\tau\neq\lambda)$，共有$\frac{1}{2}n(n-1)$个；

(ii) 3 个不同指标，$R_{\tau\lambda\mu\lambda}$，共有$\frac{1}{2}n(n-1)(n-2)$个；

(iii) 4 个不同指标，$R_{\tau\lambda\mu\nu}$，共有$\frac{1}{12}n(n-1)(n-2)(n-3)$个；

所以，协变曲率张量之不为零的独立分量总数为$\frac{1}{12}n^2(n^2-1)$个.

下面列出 $n=1-5$ 的情况：

空间维数 n：	1	2	3	4	5
独立分量数N：	0	1	6	20	50

(6) 曲率张量的缩并

曲率张量 $R^{\tau}_{\lambda\mu\nu}$ 的缩并给出

$$R_{\lambda\mu}=R^{\tau}_{\lambda\mu\tau}=\Gamma^{\tau}_{\lambda\tau,\mu}-\Gamma^{\tau}_{\lambda\mu,\tau}+\Gamma^{\sigma}_{\lambda\tau}\Gamma^{\tau}_{\sigma\mu}-\Gamma^{\sigma}_{\lambda\mu}\Gamma^{\tau}_{\sigma\tau},$$

称为 **Ricci 张量**. 它是对称张量，因为

$$\begin{aligned}R_{\lambda\mu}=R^{\tau}_{\lambda\mu\tau}&=g^{\tau\sigma}R_{\sigma\lambda\mu\tau}=g^{\tau\sigma}R_{\mu\tau\sigma\lambda}\\&=g^{\tau\sigma}R_{\tau\mu\lambda\sigma}=R^{\sigma}_{\mu\lambda\sigma}=R_{\mu\lambda}.\end{aligned}$$

还有，由于反对称性，

$$R^{\tau}_{\lambda\tau\nu}=-R^{\tau}_{\lambda\nu\tau}=-R_{\lambda\nu},$$

仍给出 Ricci 张量.

Ricci 张量 $R_{\lambda\mu}$ 再与 $g^{\lambda\mu}$ 缩乘得一标量

$$R=g^{\lambda\mu}R_{\lambda\mu},$$

称为**曲率标量**.

---

① 这里略作说明：由于对$(\tau\lambda)$的反对称关系，故在 $R_{\tau\lambda\tau\lambda}$ 和 $R_{\tau\lambda\mu\lambda}$ 二类型中有因子 1/2. 对于 $R_{\tau\lambda\mu\nu}$ 类型，由于对$(\tau\lambda)$，$(\mu\nu)$的反对称关系和$(\tau\lambda)$与$(\mu\nu)$的对称关系，出现因子$\left(\frac{1}{2}\right)^3$，而由$(\lambda\mu\nu)$之轮换关系给出因子$\left(\frac{2}{3}\right)$，故有$\left(\frac{1}{2}\right)^3\cdot\frac{2}{3}=\frac{1}{12}$.

第 6 章 二次型和主轴变换

6.1 二次型与 Hermite 型

6.1.1 二次型

定义 实数域 **R** 上具有下列形式：

$$Q = \sum_{i,j=1}^{n} k_{ij} x_i x_j$$

的表达式称为 n 个变量 $x_1, x_2, \cdots, x_n$ 的**二次型**.

不失一般性，这里将系数 k_{ij} 取为

$$k_{ij} = k_{ji} \quad (i, j = 1, 2, \cdots, n).$$

因为，若 $k'_{ij} \neq k'_{ji}$，总可以取 $k_{ij} = \frac{1}{2}(k'_{ij} + k'_{ji}) = k_{ji}$，同时由于 $x_i x_j = x_j x_i$，所以

$$k'_{ij} x_i x_j + k'_{ji} x_j x_i = (k'_{ij} + k'_{ji}) x_i x_j = (k_{ij} + k_{ji}) x_i x_j = k_{ij} x_i x_j + k_{ji} x_j x_i,$$

结果使二次型的形式不变.

另一方面，若令 $\boldsymbol{x}$ 为矢量空间 $\mathbf{V}_n(\mathbf{R})$ 中的列矢量，其转置 $\widetilde{\boldsymbol{x}}$ 为行矢量，则上述二次型还可表达为

$$Q = \sum_{i,j=1}^{n} x_i k_{ij} x_j = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} k_{ij} x_j = \widetilde{\boldsymbol{x}} \cdot K\boldsymbol{x} = (\boldsymbol{x}, K\boldsymbol{x}),$$

这里 $K = \{k_{ij}\}$ 为系数矩阵，取其为对称矩阵 $\widetilde{K} = K$，而 $(\boldsymbol{x}, K\boldsymbol{x})$ 表示标积.

6.1.2 Hermite 型

定义 对于复矢量空间 $\mathbf{V}_n(\mathbf{C})$ 中的矢量 $\boldsymbol{x}$，表达式

$$Q = \boldsymbol{x}^{\dagger} \cdot K\boldsymbol{x} = (\boldsymbol{x}, K\boldsymbol{x}) = \sum_{i,j=1}^{n} x_i^* k_{ij} x_j$$

称为 **Hermite 型**. 这里 $K = K^{\dagger}$ 是 Hermite 矩阵，上标 † 表示复共轭转置.

应该注意到，Q 总是实数，这是因为

$$Q^{\dagger} = (\boldsymbol{x}^{\dagger} \cdot K\boldsymbol{x})^{\dagger} = \boldsymbol{x}^{\dagger} K^{\dagger} \cdot \boldsymbol{x} = \boldsymbol{x}^{\dagger} \cdot K\boldsymbol{x} = Q,$$

而 Q 是标量，有 $Q^{\dagger} = Q^*$，所以 $Q = Q^*$ 为实数.

由于实数域 $\mathbf{R}_n$ 上的二次型的一切结果均可很容易地推广到复数域 $\mathbf{C}_n$ 上的

Hermite 型，所以，除在有些地方附加必要和适当的说明外，对后者一般不拟作单独讨论.

6.2 主轴变换

6.2.1 主轴变换的定义

将一个二次型(或 Hermite 型)(下面将系数矩阵表示为 $K=\{K_{ij}\}$)

$$Q \equiv \sum_{i,j} K_{ij} x_i x_j \left(\text{或 } Q \equiv \sum_{i,j} x_i^* K_{ij} x_j\right)$$

通过一个正交变换(或幺正变换)

$$\boldsymbol{x} \rightarrow \boldsymbol{y} = R\boldsymbol{x} \quad (\text{或 } \boldsymbol{x} \rightarrow \boldsymbol{y} = U\boldsymbol{x})$$

变至

$$Q = \sum_i \kappa_i y_i^2 \quad \left(\text{或 } Q = \sum_i \kappa_i \mid y_i \mid^2\right),$$

这样的变换称为**主轴变换**.

6.2.2 主轴变换的意义

(1) 几何观点

应该指出，对于二次型

$$Q = \sum_{i,j=1}^{n} K_{ij} x_i x_j = \text{常数},$$

解析几何上代表一族有心二次超曲面. 例如，

$$a x_1^2 + 2d x_1 x_2 + b x_2^2 + g = 0$$

代表有心二次曲线，又称有心圆锥曲线；而

$$a x_1^2 + b x_2^2 + c x_3^2 + 2d x_1 x_2 + 2e x_1 x_3 + 2f x_2 x_3 + g = 0$$

代表有心二次曲面.

同时，解析几何上已经知道，对于 $n=2$ 或 3，上述曲线(或曲面)族具有确定的相互垂直的主轴.

因此，预期通过坐标轴的转动(即，对任何点 P 的坐标 $\boldsymbol{x} \rightarrow \boldsymbol{y} = R\boldsymbol{x}$)可使 Q 变成

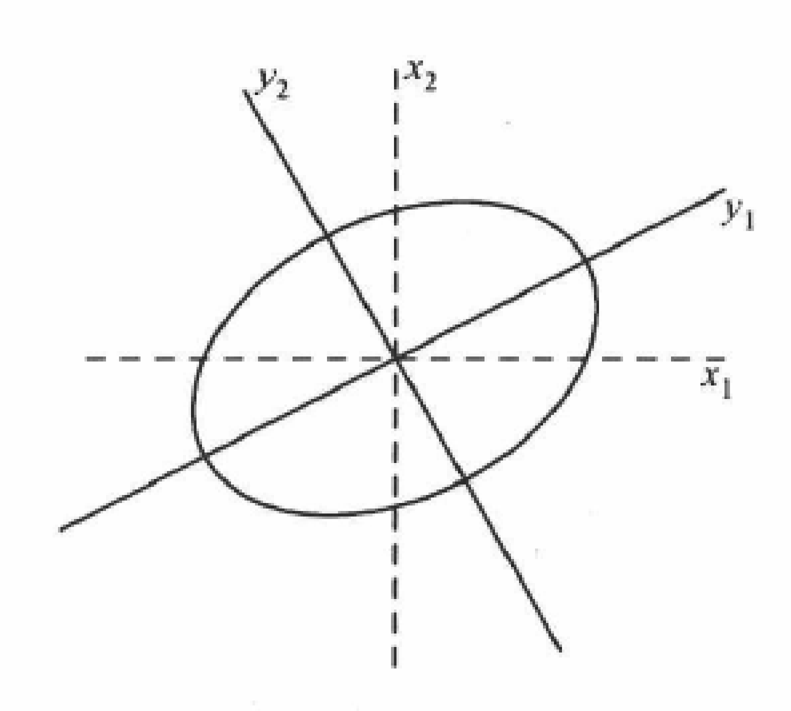
$$Q = \sum_{i=1}^{n} \kappa_i y_i^2,$$

其中 κ_i 是二次超曲面 $Q=1$ 第 i 主轴之平方的倒数，即将原曲面改为以其主轴为坐标轴，故称为**主轴变换**. 此为主轴变换之几何意义.

例如，$n=2$ 时，

$$ax_1^2+2dx_1x_2+bx_2^2=1\to\frac{y_1^2}{(1/\kappa_1)}+\frac{y_2^2}{(1/\kappa_2)}=1;$$

若 $\kappa_2>\kappa_1>0$ 时是长、短半轴分别为 $1/\sqrt{\kappa_1}$，$1/\sqrt{\kappa_2}$的椭圆，具体见下图.



(2) 代数观点

原第 j 坐标轴是 $x_j\neq0$，$x_i=0(i\neq j)$；若令 $\boldsymbol{e}^{(j)}$ 为第 j 坐标轴单位矢量：$x_j=1$，$x_i=0(i\neq j)$；则任何一点 P 以 $\boldsymbol{e}^{(1)},\cdots,\boldsymbol{e}^{(n)}$ 为基时的坐标是 $x_1,\cdots,x_n$，即

$$\boldsymbol{r}=\overrightarrow{OP}=\sum_j x_j\boldsymbol{e}^{(j)}.$$

第 i 主轴由 $y_i\neq0$，$y_j=0(j\neq i)$给出；若令 $\boldsymbol{f}^{(i)}$ 为第 i 主轴单位矢量：$y_i=1$，$y_j=0(j\neq i)$；则同一点 P 以 $\boldsymbol{f}^{(1)},\cdots,\boldsymbol{f}^{(n)}$ 为基时的坐标是 $y_1,\cdots,y_n$，即

$$\boldsymbol{r}=\overrightarrow{OP}=\sum_k y_k\boldsymbol{f}^{(k)}.$$

设以 $\boldsymbol{f}^{(1)},\cdots,\boldsymbol{f}^{(n)}$ 为基时，若 $\boldsymbol{e}^{(j)}$ 具有第 i 分量为

$$R_{ij}=(\boldsymbol{f}^{(i)},\boldsymbol{e}^{(j)})\quad(i,j=1,\cdots,n),$$

则以 $\widetilde{\boldsymbol{f}}^{(i)}$ 左乘 $\boldsymbol{r}$，由 $\widetilde{\boldsymbol{f}}^{(i)}\cdot\boldsymbol{f}^{(k)}=\delta_{ik}$ 及以上各式容易得到

$$y_i=\sum_j R_{ij}x_j.$$

另一方面，以 $\boldsymbol{e}^{(l)}$ 右乘 $\tilde{\boldsymbol{r}}$，由 $\tilde{\boldsymbol{e}}^{(j)}\cdot\boldsymbol{e}^{(l)}=\delta_{jl}$ 类似得到

$$x_l=\sum_k R_{kl}y_k.$$

由此可见，坐标变换

$$\boldsymbol{x}\to\boldsymbol{y}=R\boldsymbol{x},\quad \boldsymbol{y}\to\boldsymbol{x}=\widetilde{R}\boldsymbol{y}\Longrightarrow\widetilde{R}=R^{-1},$$

即，此为正交变换.

(3) 张量观点

对于一个变换 $\boldsymbol{x}\to\boldsymbol{y}=A\boldsymbol{x}$，有 $\boldsymbol{x}=A^{-1}\boldsymbol{y}$，于是

$$Q=\widetilde{\boldsymbol{x}}\cdot\overset{(x)}{K}\boldsymbol{x}=\widetilde{\boldsymbol{y}}\widetilde{A}^{-1}\cdot\overset{(x)}{K}A^{-1}\boldsymbol{y}=\widetilde{\boldsymbol{y}}\cdot\overset{(y)}{K}\boldsymbol{y},$$

这里 $\overset{(y)}{K}=\widetilde{A}^{-1}\overset{(x)}{K}A^{-1}$.

对于正交变换 $A=R$,有 $A^{-1}=R^{-1}=\widetilde{R}$,所以

$$\overset{(y)}{K} = R\overset{(x)}{K}R^{-1} = R\overset{(x)}{K}\widetilde{R},$$

或者

$$\overset{(y)}{K}_{ik} = \sum_{j,l} R_{ij}\overset{(x)}{K}_{jl}R_{kl} = \sum_{j,l} R_{ij}R_{kl}\overset{(x)}{K}_{jl},$$

这正是 Descartes 张量的变换关系.

(4) 主轴变换看作本征值问题

所以主轴变换就是要寻求这样的一个正交矩阵 R 使

$$RKR^{-1} = D = \mathrm{diag}[\kappa_1, \kappa_2, \cdots, \kappa_n],$$

即 D 是对角元为 $\kappa_1, \cdots, \kappa_n$ 的对角矩阵.

称 $R^{-1}=\widetilde{R}=S(S_{ij}=R_{ji})$,于是 $S^{-1}KS=D$,即

$$KS = SD \quad 或 \quad \sum_j K_{ij}S_{jk} = S_{ik}\kappa_k \quad (i = 1, \cdots, n);$$

因此,化为求解下列**本征值问题**:

$$\sum_j K_{ij}\xi_j = \kappa\xi_i \quad 或 \quad K\boldsymbol{\xi} = \kappa\boldsymbol{\xi},$$

其第 k 个**本征值为** κ_k,相应**本征矢** $\boldsymbol{\xi}^{(k)}=(S_{ik})$ 即为 $\boldsymbol{f}^{(k)}$.

可以认为,$K=\{K_{ij}\}$是一算符,当作用在矢量 $\boldsymbol{\xi}$ 上时,将其变为 $\boldsymbol{\xi}^K$;当其作用在平行于主轴之矢量上时,并不改变其方向.

下节来讨论一般的本征值问题.

6.3 本征值问题①

设 $A=\{a_{ij}\}$为域 $\mathbf{F}$($\mathbf{R}$ 或 $\mathbf{C}$)上的 n 阶矩阵. $\boldsymbol{x}$ 为矢量空间$\mathbf{V}_n(\mathbf{F})$中的列矢量,λ 为一标量. 若变换 $A\boldsymbol{x}$ 给出 $\lambda\boldsymbol{x}$,即

$$A\boldsymbol{x} = \lambda\boldsymbol{x},$$

则此 $\boldsymbol{x}$ 称为矩阵 A 的**本征矢**,而 λ 称为矩阵 A 的与此本征矢 $\boldsymbol{x}$ 对应的**本征值**. 求解矩阵的本征矢和本征值的问题则称为**本征值问题**.

6.3.1 本征值的确定及其性质

上述方程还可写成

$$(A - \lambda I)\boldsymbol{x} = 0,$$

若 $\widetilde{\boldsymbol{x}}=[x_1 \quad x_2 \quad \cdots \quad x_n]$,上式是关于 $x_1, \cdots, x_n$ 的齐次线性方程组,有非平凡解 $\boldsymbol{x}\neq$

① 参考:A. W. Joshi, *Matrices and Tensors in Physics*, pp. 88—101.

0 的条件是，当且仅当

$$D(\lambda) \equiv \det(A-\lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0.$$

显然，行列式 $D(\lambda)$是 λ 的 n 次多项式，称为矩阵 A 的**本征多项式**；而 $D(\lambda)=0$ 则称为 A 的**本征方程**. 本征方程的 n 个根 $\lambda_1,\cdots,\lambda_n$ 即为矩阵的 n 个本征值. 于是，有下列结果：

$$D(\lambda) = \prod_{i=1}^{n}(\lambda_i-\lambda),$$

$$\sum_{i=1}^{n}\lambda_i = \sum_{i=1}^{n}a_{ii} = \mathrm{tr}A,$$

$$\prod_{i=1}^{n}\lambda_i = \det A.$$

矩阵的本征值集合称为其**本征值谱**.

同时可以证明，幺正矩阵 U 的全体本征值，其数值均为 1. 这是因为

$$U\boldsymbol{x} = \lambda\boldsymbol{x}, \quad \boldsymbol{x}^{\dagger}U^{\dagger} = \lambda^{*}\boldsymbol{x}^{\dagger} \Longrightarrow \boldsymbol{x}^{\dagger}U^{\dagger}U\boldsymbol{x} = \lambda\lambda^{*}\boldsymbol{x}^{\dagger}\boldsymbol{x};$$

由于 $U^{\dagger}U=I$，所以 $\boldsymbol{x}^{\dagger}\boldsymbol{x}(|\lambda|^2-1)=0$，对于 $\boldsymbol{x}\neq 0$ 有 $|\lambda|^2=1$ 或 $|\lambda|=1$.

另外还可证明，对于 Hermite 矩阵(或对称矩阵)K，其全体本征值均为实数. 这是因为

$$U\boldsymbol{x} = \lambda\boldsymbol{x}, \quad \boldsymbol{x}^{\dagger}U^{\dagger} = \lambda^{*}\boldsymbol{x}^{\dagger} \Longrightarrow \boldsymbol{x}^{\dagger}U\boldsymbol{x} = \lambda\boldsymbol{x}^{\dagger}\boldsymbol{x},$$

$$\boldsymbol{x}^{\dagger}U^{\dagger}\boldsymbol{x} = \lambda^{*}\boldsymbol{x}^{\dagger}\boldsymbol{x};$$

由于 $U^{\dagger}=U$，所以$(\lambda-\lambda^{*})\boldsymbol{x}^{\dagger}\cdot\boldsymbol{x}=0$；对于 $\boldsymbol{x}\neq 0$ 有 $\lambda^{*}=\lambda$，故 λ 为实数. 其实，因为

$$K\boldsymbol{x} = \lambda\boldsymbol{x} \Longrightarrow \boldsymbol{x}^{\dagger}\cdot K\boldsymbol{x} = \lambda\boldsymbol{x}^{\dagger}\cdot\boldsymbol{x},$$

由于标积 $\boldsymbol{x}^{\dagger}\cdot\boldsymbol{x}$ 和 Hermite 型 $\boldsymbol{x}^{\dagger}K\boldsymbol{x}$ 均为实数，故 λ 为实数. [只要将 † 换成 ^{T}，即可证明对称矩阵的情况.]

n 个本征值不一定全是相异的，因为 $D(\lambda)=0$ 可能有重根. 若本征值 λ_i 的**重数**为 m_i，而相异本征值为 s 个$(i=1,\cdots,s)$，即 $D(\lambda)=\prod_{i=1}^{s}(\lambda_i-\lambda)^{m_i}$，显然有 $\sum_{i=1}^{s}m_i=n$.

6.3.2 本征矢及其性质・矩阵的对角化

与 n 个本征值相对应的 n 个本征矢不一定全是相互无关的. 关于本征矢的性质，现在来介绍以下一些结果.

(1) 相异本征值的情况

与相异本征值对应的本征矢是相互无关的.

证 令 $\lambda_1,\cdots,\lambda_s$ 为 $s(\leqslant n)$ 个相异本征值，而 $\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(s)}$ 为与之对应的本征矢. 假定有数 $\alpha_1,\cdots,\alpha_s$ 使

$$\sum_{i=1}^{s}\alpha_i\boldsymbol{x}^{(i)} = 0,$$

要证明 $\alpha_i=0(i=1,\cdots,s)$. 首先对上式两边左乘以 $(\lambda_1 I-A)$，注意到 $(\lambda_1 I-A)\boldsymbol{x}^{(1)}=0$ 和 $A\boldsymbol{x}^{(i)}=\lambda_i\boldsymbol{x}^{(i)}(i=2,\cdots,s)$，于是

$$\sum_{i=2}^{s}(\lambda_1-\lambda_i)\alpha_i\boldsymbol{x}^{(i)} = 0.$$

对此式再左乘以 $(\lambda_2 I-A)$ 后变成

$$\sum_{i=3}^{s}(\lambda_1-\lambda_i)(\lambda_2-\lambda_i)\alpha_i\boldsymbol{x}^{(i)} = 0.$$

经过 $s-1$ 次这样的运算后得到

$$(\lambda_1-\lambda_s)(\lambda_2-\lambda_s)\cdots(\lambda_{s-1}-\lambda_s)\alpha_s\boldsymbol{x}^{(s)} = 0;$$

这意味着 $\alpha_s=0$. 但本征值和本征矢的顺序是任意的，同样可证明 $\alpha_1=\alpha_2=\cdots=\alpha_{s-1}=0$. 因此，本征矢 $\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(s)}$ 是线性无关的.

(2) 多重本征值的情况

(i) n 阶矩阵 A 的本征值 λ_i，重数为 $m_i(i=1,\cdots,s)$ 时，若

$$\mathrm{rank}(\lambda_i I-A) = n-m_i,$$

则与 λ_i 对应有 m_i 个线性无关本征矢，称为**简并本征矢**.（与 $m_i=1$ 对应的称为**非简并本征矢**.）

证 确定本征矢的方程

$$(A-\lambda_i I)\boldsymbol{x}^{(i)} = 0,$$

即

$$\sum_{s=1}^{n}(a_{rs}-\lambda_i\delta_{rs})x_s^{(i)} = 0 \quad (r=1,\cdots,n),$$

其中只有 $(n-m_i)$ 个是独立的；因此，比如说，可将 $x_{m_i+1}^{(1)},\cdots,x_n^{(i)}$ 解出为 $x_1^{(i)},\cdots,x_{m_i}^{(i)}$ 的线性函数，它具有 m_i 重任意性，所以总可选择 m_i 个线性无关本征矢. 例如，前 m_i 个分量的下列 m_i 种选择：$(1,0,\cdots,0)^{\mathrm{T}},(0,1,\cdots,0)^{\mathrm{T}},\cdots,(0,0,\cdots,1)^{\mathrm{T}}$ 就是如此.

(ii) 容易证明 $n-m_i\leqslant\mathrm{rank}(\lambda_i I-A)\leqslant n-1$.

因为 $\det(\lambda_i I-A)=-D(\lambda)=0$，故其秩 $<n$，右端得证. 若左端为 $<n-m_i$ 会导致与 λ_i 对应的线性无关本征矢的个数 $>m_i$，而使矩阵 A 的线性无关本征矢的总数

会 $> \sum_{i=1}^{s} m_i = n$，显然不正确；左端得证.

(3) 非简并本征矢互相正交

对于 Hermite 矩阵(或对称矩阵)，对应于相异本征值的本征矢互相正交.

证 令 $\boldsymbol{\xi}_1$ 和 $\boldsymbol{\xi}_2$ 为 Hermite 矩阵 K 的分别与任意二相异本征值 λ_1 和 λ_2 对应的两个本征矢，即

$$K\boldsymbol{\xi}_1 = \lambda_1 \boldsymbol{\xi}_1, \quad K\boldsymbol{\xi}_2 = \lambda_2 \boldsymbol{\xi}_2 \rightarrow \boldsymbol{\xi}_2^{\dagger} K = \lambda_2 \boldsymbol{\xi}_2^{\dagger},$$

后者用了 $K^{\dagger}=K$ 和 $\lambda_2^*=\lambda_2$，分别左乘以 $\boldsymbol{\xi}_2^{\dagger}$ 和右乘以 $\boldsymbol{\xi}_1$ 并相减得

$$(\lambda_1 - \lambda_2)\boldsymbol{\xi}_2^{\dagger} \cdot \boldsymbol{\xi}_1 = 0 \rightarrow \boldsymbol{\xi}_2^{\dagger} \cdot \boldsymbol{\xi}_1 = 0.$$

得证.

(4) 简并本征矢可选择为正交

同一本征值的不同本征矢，可以选择为互相正交.

证 设已给 $A\boldsymbol{x}^{(k)} = \lambda \boldsymbol{x}^{(k)} (k=1,2,\cdots,m)$，$\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(m)}$ 不正交，可应用 Schmidt 正交化法或正交归一化法选出互相正交的本征矢.

现以 $m=3$ 为例予以阐明.

正交化法选择正交矢 $\boldsymbol{f}^{(1)},\boldsymbol{f}^{(2)},\boldsymbol{f}^{(3)}$ 如下：

$$\boldsymbol{f}^{(1)} = \boldsymbol{x}^{(1)}, \quad \boldsymbol{f}^{(2)} = \boldsymbol{x}^{(2)} - \boldsymbol{x}^{(1)} \frac{(\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)})}{(\boldsymbol{x}^{(1)},\boldsymbol{x}^{(1)})},$$

则

$$(\boldsymbol{f}^{(1)},\boldsymbol{f}^{(2)}) \equiv 0;$$

$$\boldsymbol{f}^{(3)} = \boldsymbol{x}^{(3)} - \boldsymbol{f}^{(1)} \frac{(\boldsymbol{f}^{(1)},\boldsymbol{x}^{(3)})}{(\boldsymbol{f}^{(1)},\boldsymbol{f}^{(1)})} - \boldsymbol{f}^{(2)} \frac{(\boldsymbol{f}^{(2)},\boldsymbol{x}^{(3)})}{(\boldsymbol{f}^{(2)},\boldsymbol{f}^{(2)})},$$

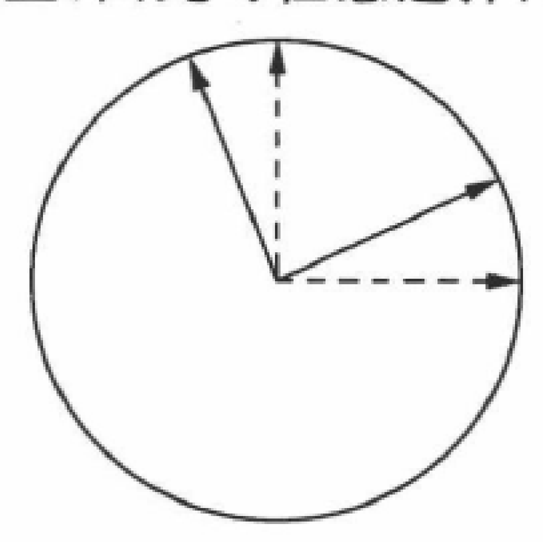
则

$$(\boldsymbol{f}^{(1)},\boldsymbol{f}^{(3)}) \equiv 0, \quad (\boldsymbol{f}^{(2)},\boldsymbol{f}^{(3)}) \equiv 0.$$

正交归一化法选择正交归一矢 $\boldsymbol{e}^{(1)},\boldsymbol{e}^{(2)},\boldsymbol{e}^{(3)}$ 如下：

$$\boldsymbol{e}^{(1)} = \frac{\boldsymbol{f}^{(1)}}{(\boldsymbol{f}^{(1)},\boldsymbol{f}^{(1)})}, \quad \boldsymbol{e}^{(2)} = \frac{\boldsymbol{f}^{(2)}}{(\boldsymbol{f}^{(2)},\boldsymbol{f}^{(2)})}, \quad \boldsymbol{e}^{(3)} = \frac{\boldsymbol{f}^{(3)}}{(\boldsymbol{f}^{(3)},\boldsymbol{f}^{(3)})}.$$

当然，这样的选择不唯一. 以 $m=2$ 为例，这在几何上意味着两个本征矢，即两个主轴都在一截面为圆的平面上，因此可任意选择两个互相垂直的主轴；如下图.



(5) 对称矩阵的本征矢可选择为实矢量.

证 $$K\boldsymbol{\xi}=\kappa\boldsymbol{\xi},\quad K\boldsymbol{\xi}^*=\kappa\boldsymbol{\xi}^*\quad\Longrightarrow\quad K\frac{\boldsymbol{\xi}+\boldsymbol{\xi}^*}{2}=\kappa\frac{\boldsymbol{\xi}+\boldsymbol{\xi}^*}{2},$$

因为 K 为实矩阵导致 κ 为实数,故可选本征矢 $\frac{1}{2}(\boldsymbol{\xi}+\boldsymbol{\xi}^*)$ 为实矢量.

(6) 矩阵的对角化

迄今已知,n 阶 Hermite 矩阵 K 具有 n 个实本征值 $\kappa_1,\cdots,\kappa_n$ 和 n 个线性无关并正交的本征矢 $\boldsymbol{\xi}^{(1)},\cdots,\boldsymbol{\xi}^{(n)}$. 现在进一步设 $\kappa_1\geqslant\kappa_2\geqslant\cdots\geqslant\kappa_n$ 以及 $\boldsymbol{\xi}^{(r)\dagger}\cdot\boldsymbol{\xi}^{(s)}=\delta_{rs}$,因为实数总可排序,而本征矢的归一化仍为本征矢.

用 $\boldsymbol{\xi}^{(1)},\cdots,\boldsymbol{\xi}^{(n)}$ 构造矩阵

$$\varXi=[\boldsymbol{\xi}^{(1)},\boldsymbol{\xi}^{(2)},\cdots,\boldsymbol{\xi}^{(n)}],$$

显然,$\varXi^\dagger\varXi=I_n$,它是幺正矩阵. 于是,由 $K\boldsymbol{\xi}^{(r)}=\kappa_r\boldsymbol{\xi}^{(r)}$ 可证实

$$\varXi^\dagger K\varXi=[\boldsymbol{\xi}^{(1)},\boldsymbol{\xi}^{(2)},\cdots,\boldsymbol{\xi}^{(n)}]^\dagger[\kappa_1\boldsymbol{\xi}^{(1)},\kappa_2\boldsymbol{\xi}^{(2)},\cdots,\kappa^{(n)}\boldsymbol{\xi}^{(n)}]=\mathrm{diag}[\kappa_1,\kappa_2,\cdots,\kappa_n],$$

即 Hermite 矩阵 K 可通过幺正矩阵使之对角化.

同时注意到 $\varXi^\dagger=\varXi^{-1}$,故称 K 与 $\mathrm{diag}[\kappa_1\quad\kappa_2\quad\cdots\quad\kappa_n]$ **幺正相似**.

另一方面,Hermite 型 $Q=\boldsymbol{x}^\dagger K\boldsymbol{x}$ 通过幺正变换 $\boldsymbol{x}=\varXi\boldsymbol{\xi}$ 得

$$\begin{aligned}Q&=\boldsymbol{x}^\dagger K\boldsymbol{x}=\boldsymbol{\xi}^\dagger\varXi^\dagger\cdot K\varXi\boldsymbol{\xi}\\&=\boldsymbol{\xi}^\dagger\cdot\mathrm{diag}[\kappa_1\quad\kappa_2\quad\cdots\quad\kappa_n]\boldsymbol{\xi}=\sum_{i=1}^n\kappa_i\,|\xi_i|^2,\end{aligned}$$

即此幺正变换就是主轴变换;而此形式称为 Hermite 型在幺正变换群下的**第二标准型**.

上述这些结果常称为**主轴定理**.

因此,若能找到幺正矩阵 $\varXi$ 能使矩阵 K 对角化,则对角元即其本征值,而 $\varXi$ 的列矢量即其相应本征矢. 这提供了求解问题的一种方法.

(7) 可对易 Hermite 矩阵的本征矢

设 A,B 为 n 阶可对易 Hermite 矩阵

$$AB-BA=0.$$

首先,由 A 的本征值方程 $A\boldsymbol{x}=\lambda\boldsymbol{x}$ 可得

$$BA\boldsymbol{x}=A(B\boldsymbol{x})=\lambda(B\boldsymbol{x}).$$

这就是说,若 $\boldsymbol{x}$ 为与 A 的本征值 λ 对应的本征矢,则 $B\boldsymbol{x}$ 也是具有同一本征值 λ 的本征矢. 对于非简并本征矢,必然有

$$B\boldsymbol{x}=\mu\boldsymbol{x},$$

即 $\boldsymbol{x}$ 同时也是矩阵 B 具有本征值为 μ 的本征矢.

对于 A 的本征值 λ 有 m 个简并本征矢 $\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(m)}$ 的情况,则一般有

$$BX=XC,$$

这里 $X=[\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(m)}]$是 $n\times m$ 矩阵,$C=\{c_{ij}\}$是 m 阶方阵. 现在用 $X^{\dagger}$ 左乘上式两边,假定简并本征矢 $\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(m)}$已正交归一化,则有 $X^{\dagger}X=I_m$;因而

$$X^{\dagger}BX = X^{\dagger}XC = C,$$

$$C^{\dagger} = (X^{\dagger}BX)^{\dagger} = X^{\dagger}B^{\dagger}X = X^{\dagger}BX = C;$$

即 C 是 m 阶 Hermite 矩阵,可以对角化. 若有非奇异 m 阶矩阵 $D=\{d_{rs}\}$使 C 对角化,

$$D^{-1}CD = \mathrm{diag}[\mu_1,\mu_2,\cdots,\mu_m];$$

作变换

$$Y = XD \quad 即 \quad X = YD^{-1},$$

这里 $Y=[\boldsymbol{y}^{(1)},\boldsymbol{y}^{(2)},\cdots,\boldsymbol{y}^{(m)}]$是 $n\times m$ 矩阵,则有

$$BY = BXD = XCD = XDD^{-1}CD = Y\mathrm{diag}[\mu_1,\cdots,\mu_m],$$

即(对 $1\leqslant r\leqslant m$)

$$\boldsymbol{y}^{(r)} = \sum_{s=1}^{m}\boldsymbol{x}^{(s)}d_{sr},$$

$$B\boldsymbol{y}^{(r)} = \mu_r\boldsymbol{y}^{(r)},$$

$$A\boldsymbol{y}^{(r)} = \lambda\boldsymbol{y}^{(r)};$$

也就是说,$\boldsymbol{y}^{(1)},\cdots,\boldsymbol{y}^{(m)}$是 A 和 B 的共同本征矢.

对 A 的每个不同本征值继续执行此过程,最后可求得 A 和 B 的全部共同本征矢.

总结起来,得到下列在矩阵代数中以及量子力学中都很重要的**定理**:对于相互可对易 Hermite 矩阵,有可能求得其共同本征矢集合. 或者换句话说,**全体相互可对易 Hermite 矩阵可同时对角化.**

6.4 本征值的极值性质

还可利用极值原理对二次型(Hermite 型)

$$Q(\boldsymbol{x}) = (\boldsymbol{x},K\boldsymbol{x})$$

作主轴变换. 为此,先引进下列定理.

6.4.1 极值原理

(1) Weierstrass 定理

限定在有限闭域内的多元连续函数,在该区域内某处达到其上(或下)极限.

(2) 极值原理

由于正交变换(幺正变换)下标积不变,即对这样的变换 $\boldsymbol{x}\to\boldsymbol{y}$ 有

$$Q(\boldsymbol{x}) = Q(\boldsymbol{y}) \quad 及 \quad (\boldsymbol{x},\boldsymbol{x}) = (\boldsymbol{y},\boldsymbol{y}).$$

若变换能使 $Q(\boldsymbol{y})$ 变为平方项之和,即

$$Q(\boldsymbol{y}) = \sum \kappa_i y_i^2 \leqslant \kappa_1(\boldsymbol{y},\boldsymbol{y})$$

(对 Hermite 型,y_i^2 应换为$|y_i|^2$),这里 κ_1 为 κ_i 中最大者,此即主轴变换.若令

$$\kappa_1 \geqslant \kappa_2 \geqslant \cdots \geqslant \kappa_n,$$

则显然有

$$\kappa_n(\boldsymbol{x},\boldsymbol{x}) \leqslant Q(\boldsymbol{x}) \leqslant \kappa_1(\boldsymbol{x},\boldsymbol{x});$$

此即极值原理之依据.

6.4.2 主轴变换的具体步骤

为方便起见,下面具体讨论二次型(Hermite 型完全类似).

首先,限制 $\boldsymbol{x}=(x_1,\cdots,x_n)$在单位球面上变化,即限制条件为

$$(\boldsymbol{x},\boldsymbol{x}) = \sum x_i^2 = 1.$$

于是,根据 Weierstrass 定理,在球面上某点(比如说)

$$x_i = R_{1i} \quad (i = 1,\cdots,n)$$

(即 $\boldsymbol{x}$ 的基中矢量 $\boldsymbol{f}^{(1)}$之坐标,或$(\boldsymbol{f}^{(1)}\cdot\boldsymbol{e}^{(i)})$),$Q(\boldsymbol{x})$达到其上极限(比如说)$\kappa_1$.

其次,考虑(单位球面) $\sum x_i^2 = 1$ 和(与 $\boldsymbol{f}^{(1)}$ 垂直之平面) $\sum R_{1i}x_i = 0$ 相交的截面.根据 Weierstrass 定理,在此子区域上某点(比如说)

$$x_i = R_{2i} \quad (i = 1,\cdots,n)$$

(即 $\boldsymbol{x}$ 的基中矢量 $\boldsymbol{f}^{(2)}$之坐标,或$(\boldsymbol{f}^{(2)}\cdot\boldsymbol{e}^{(i)})$);$Q(\boldsymbol{x})$达到其上极限(比如说)$\kappa_2\leqslant\kappa_1$.因此,$\boldsymbol{f}^{(1)}\perp\boldsymbol{f}^{(2)}$,两个单位矢量相互垂直.

依此继续进行下去,直至求得

$$\kappa_1 \geqslant \kappa_2 \geqslant \cdots \geqslant \kappa_n,$$
$$\boldsymbol{f}^{(1)},\boldsymbol{f}^{(2)},\cdots,\boldsymbol{f}^{(n)};$$

后者构成一个新基.

利用此新基,则作为 P 点的新坐标 y_i 与原坐标 x_j 之间的关系为

$$y_i = \sum_{j=1}^{n} R_{ij}x_j,$$

只要注意到对于第 k 个 $\boldsymbol{f}$ 矢量 $\boldsymbol{f}^{(k)}$ 有

$$x_j = R_{kj}, \quad 所以 \quad y_i = \delta_{ik},$$

上式不难验证.

现在,$Q(y)$的形式必然为 $\tilde{y}\overset{(y)}{K}y$. 因为对于 $y_1=1,y_2=\cdots=y_n=0$ 时,其值是 κ_1,所以$\overset{(y)}{K}_{11}=\kappa_1$. 这个值 κ_1 为单位球面 $1=\sum x_i^2=\sum y_i^2$ 上的最大值,所以

$$\frac{\tilde{y}}{\sqrt{\sum y_i^2}}\overset{(y)}{K}\frac{y}{\sqrt{\sum y_i^2}}-\kappa_1\leqslant 0 \quad 或 \quad \tilde{y}\overset{(y)}{K}y-\kappa_1\sum y_i^2\leqslant 0.$$

由此得出下列结果:

$$\overset{(y)}{K}_{12}=0,\cdots, \quad 或 \quad \overset{(y)}{K}_{1i}=0 \quad (i=2,\cdots,n).$$

这可证明如下:令

$$y_1=1-O(\varepsilon^2),\ y_2=\varepsilon,\ y_3=\cdots=y_n=0,$$

则前式变为要求

$$\tilde{y}\overset{(y)}{K}y-\kappa_1\sum y_i^2=2\overset{(y)}{K}_{12}\varepsilon+O(\varepsilon^2)\leqslant 0,$$

但若$\overset{(y)}{K}_{12}\neq 0$,则当 $\varepsilon\to\pm 0$ 时,总可对 ε 选择适当正负号使左边为正而不满足,故要求$\overset{(y)}{K}_{12}=0$,同理可证明$\overset{(y)}{K}_{1i}=0\ (i\neq 1)$. 于是

$$Q=\kappa_1 y_1^2+Q'(y_2,\cdots,y_n),$$

对 Q'可类似地进行论证得出$\overset{(y)}{K}_{2i}=0\ (i\neq 2)$. 继续进行下去,直至最后得到

$$\overset{(y)}{K}_{ij}=\kappa_i\delta_{ij},$$

$$Q=\sum_{i,j}y_i\overset{(y)}{K}_{ij}y_j=\sum_{i=1}^{n}\kappa_i y_i^2.$$

6.4.3 变分形式

综上所述,还可写成下列变分形式:

$$\max_{|x|>0}\frac{\tilde{x}Kx}{\tilde{x}x}=\kappa_1, \quad \min_{|x|>0}\frac{\tilde{x}Kx}{\tilde{x}x}=\kappa_n,$$

或

$$\max_{|x|=1}\tilde{x}Kx=\kappa_1, \quad \min_{|x|=1}\tilde{x}Kx=\kappa_n.$$

进一步还可有下列表述(max 下为约束条件):

$$\max_{|x|>0}\frac{\tilde{x}Kx}{\tilde{x}x}=\kappa_1, \quad 在\ x=f^{(1)}\ 时达到;$$

$$\max_{\substack{|x|>0\\(f^{(1)},x)=0}}\frac{\tilde{x}Kx}{\tilde{x}x}=\kappa_2, \quad 在\ x=f^{(2)}\ 时达到;$$

$$\max_{\substack{|x|>0\\(f^{(i)},x)=0\\(i=1,\cdots,r-1)}}\frac{\tilde{x}Kx}{\tilde{x}x}=\kappa_r, \quad 在\ x=f^{(r)}\ 时达到.$$

注意,对于 Hermite 型,只需将 $\tilde{x}$ 换成 $x^\dagger$ 即可.另外,$\dfrac{(x,Kx)}{(x,x)}$有时称为 **Rayleigh-Ritz 比**.

6.5 Sylvester 惯性律

已知主轴变换将二次型(Hermite 型)化为平方项之和,其系数即其本征值为实数.若 n 个本征值中前 p 个为正值,其次 $r-p$ 个为负值,其余 $n-r$ 个为零;若作代换

$$\zeta_i = (\kappa_i)^{1/2}\xi_i \quad (i = 1,\cdots,p),$$
$$\zeta_j = (-\kappa_j)^{1/2}\xi_j \quad (j = p+1,\cdots,r);$$

则有(对 Hermite 型,下式中的 $\zeta^2 \to \zeta^* \zeta = |\zeta|^2$)

$$Q = \sum_{i=1}^{p}\zeta_i^2 - \sum_{j=p+1}^{r}\zeta_j^2 = \zeta_1^2 + \cdots + \zeta_p^2 - \zeta_{p+1}^2 - \cdots - \zeta_r^2.$$

约化形式中正(和负)平方项个数 p(和 $r-p$)对给定函数 Q 是不变量,不依赖于归约方法(**Sylvester 惯性律**).

证 假定有另一约化形式

$$Q = \eta_1^2 + \cdots + \eta_q^2 - \eta_{q+1}^2 - \cdots - \eta_r^2,$$

其中有 $q(\neq p)$个正项(和$(r-q)$个负项).因为两者都是由同一$Q(\boldsymbol{x})$通过非奇异线性变换(坐标变换)而获得,两者之间必有一非奇异线性变换相联系.若 $q<p$:

$\boldsymbol{x}\to\boldsymbol{\zeta}$:$Q(x)\geqslant 0$,当 $\zeta_{p+1}=\cdots=\zeta_r=0$,即在 $n-(r-p)$维子空间 S_1 中;

$\boldsymbol{x}\to\boldsymbol{\eta}$:$Q(x)<0(\boldsymbol{x}\neq\boldsymbol{0})$,当 $\eta_1=\cdots=\eta_q=\eta_{r+1}=\cdots=\eta_n=0$,即在 $r-q$ 维子空间 S_2 中.由于 S_1 和 S_2 的维数和

$$n-(r-p)+r-q = n+p-q > n,$$

故在其交空间 $S_1\cap S_2$ 会出现矛盾.若 $q>p$,同样会推出矛盾.故 $q=p$,得证.

约化形式

$$Q = \sum_{i=1}^{p}\xi_i^2 - \sum_{j=p+1}^{r}\xi_j^2$$

称为二次型在一般线性群下的**(第一)标准型**.

正负项个数之差称为**符号差**

$$\mathrm{sig}Q = p-(r-p) = 2p-r,$$

r 为二次型的**秩**.

习 题

6.1 求矩阵 $A=\begin{bmatrix} q & p & p \\ p & q & p \\ p & p & q \end{bmatrix}$ 的本征值和本征矢，其中 p 和 q 是标量，$p\neq 0$.

解 容易求得本征方程为

$$\det(A-\lambda I)=D(\lambda)=(q-p-\lambda)^2(q+2p-\lambda)=0,$$

因而本征值为 $\lambda_1=\lambda_2=q-p$，$\lambda_3=q+2p$.

对于 $\lambda_3=q+2p$，其相应本征矢由 $(A-\lambda_3 I)\boldsymbol{x}=\boldsymbol{0}$ 给出；因为 $\text{rank}(A-\lambda_3 I)$ 为 2，可求得 $x_3=x_2=x_1$，x_1 任意，故可得

$$\boldsymbol{x}^{(3)}=(c\ c\ c)^{\mathrm{T}} \quad 或 \quad (1\ \ 1\ \ 1)^{\mathrm{T}} \quad 或 \quad \frac{1}{\sqrt{3}}(1\ \ 1\ \ 1)^{\mathrm{T}}.$$

对于 $\lambda_1=\lambda_2=q-p$，其相应本征矢由 $(A-\lambda_1 I)\boldsymbol{x}=\boldsymbol{0}$ 给出；其矩阵之秩 $r=1=n-m$，可求得 $x_3=-x_1-x_2$，x_1 和 x_2 任意，故一般形式为 $(a\ b\ \ -a-b)^{\mathrm{T}}$，可取

$$\boldsymbol{x}^{(1)}=(a\ \ 0\ \ -a)^{\mathrm{T}} \quad 或 \quad (1\ \ 0\ \ -1)^{\mathrm{T}} \quad 或 \quad \frac{1}{\sqrt{2}}(1\ \ 0\ \ -1)^{\mathrm{T}},$$

$$\boldsymbol{x}^{(2)}=(0\ \ b\ \ -b)^{\mathrm{T}} \quad 或 \quad (0\ \ 1\ \ -1)^{\mathrm{T}} \quad 或 \quad \frac{1}{\sqrt{2}}(0\ \ 1\ \ -1)^{\mathrm{T}}.$$

但 $\widetilde{\boldsymbol{x}}^{(1)}\cdot\boldsymbol{x}^{(2)}=ab$(或 1，或 1/2)$\neq 0$，可改为选 $\boldsymbol{x}^{(2)}=\frac{1}{\sqrt{6}}(1\ \ -2\ \ 1)^{\mathrm{T}}$ 使之正交.

所以，三个相互无关的正交归一本征矢为

$$\boldsymbol{x}^{(1)}=\frac{1}{\sqrt{2}}(1\ \ 0\ \ -1)^{\mathrm{T}},\quad \boldsymbol{x}^{(2)}=\frac{1}{\sqrt{6}}(1\ \ -2\ \ 1)^{\mathrm{T}},\quad \boldsymbol{x}^{(3)}=\frac{1}{\sqrt{3}}(1\ \ 1\ \ 1)^{\mathrm{T}}.$$

6.2 求下列各矩阵的本征值和本征矢：

(1) $\begin{bmatrix} 1 & 0 & -2 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$, (2) $\begin{bmatrix} 1 & 1 & -2 \\ 1 & 2 & -1 \\ -2 & -1 & 1 \end{bmatrix}$, (3) $\begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix}$,

(4) $\begin{bmatrix} 5 & -2 & -4 \\ -2 & 2 & 2 \\ -4 & 2 & 5 \end{bmatrix}$, (5) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, (6) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

(7) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

解 前为本征值，后为本征矢；n 为矩阵之阶，m 为本征值之重数，$r=\text{rank}(\lambda I$

$-A)$，$i^2=-1$.

(1) $0,2+i,2-i;[-2,2,-1]^T,[1-i,-i,-1]^T,[1+i,i,-1]^T$.

(2) $4,1,-1;\frac{1}{\sqrt{6}}(\sqrt{2},\sqrt{2},\sqrt{2})^T,\frac{1}{\sqrt{6}}(-1,2,1)^T,\frac{1}{\sqrt{6}}(\sqrt{3},0,\sqrt{3})^T$.

(3) $9,3,-3;\frac{1}{3}(-1,2,2)^T,\frac{1}{3}(2,2,-1)^T,\frac{1}{3}(2,-1,2)^T$.

(4) $10,1,1;\frac{1}{3}(2,-1,-2)^T,\frac{1}{3\sqrt{2}}(1,4,-1)^T,\frac{1}{\sqrt{2}}(-1,0,-1)^T$.（对 $\lambda=1$ 而言，$n=3,m=2,r=1,r=n-m$.）

(5) $1,\frac{1}{2}(-1+i\sqrt{3}),\frac{1}{2}(-1-i\sqrt{3});(1,1,1)^T,(-1+i\sqrt{3},-1-i\sqrt{3},2)^T$, $(-1-i\sqrt{3},-1+i\sqrt{3},2)^T$.

(6) $1,1,1;(1,0,0)^T,(0,1,0)^T,(0,0,1)^T$.（$n=3,m=3,r=0;r=n-m$.）

(7) $1,1,1;x^{(1)}=x^{(2)}=x^{(3)}=(1,0,0)^T$.（$n=3,m=3,r=2;r>n-m$.）

因为(1)和(5)是非对称矩阵，有复本征值和本征矢. 因为(7)中 $r>n-m$，本征矢线性相关. 在(2)，(3)，(4)和(6)中的本征矢是线性无关且正交归一的.

6.3 求二次型 $Q=3(x_1^2+x_2^2+x_3^2)+2(x_1x_2+x_2x_3+x_3x_1)$ 的秩和惯性，并确定使 Q 化为标准型的坐标变换.

解 与此 Q 对应的对称矩阵 K 是

$$K=\begin{bmatrix}3&1&1\\1&3&1\\1&1&3\end{bmatrix},$$

可求得本征值为 $\lambda_1=5,\lambda_2=\lambda_3=2$；于是 Q 的第二和第一标准型是

$$Q=\sum_{i=1}^{r}\lambda_i y_i^2=5y_1^2+2y_2^2+2y_3^2$$

$$=\sum_{i=1}^{p}z_i^2-\sum_{j=1}^{r-p}z_j^2=z_1^2+z_2^2+z_3^2;$$

因此秩 $r=3$，惯性 $\mathrm{In}Q=\mathrm{In}K=\{p,r-p,n-r\}=\{3,0,0\}$.

为确定正交变换，先求出其正交归一本征矢为

$$\frac{1}{\sqrt{3}}(1,1,1)^T,\frac{1}{\sqrt{2}}(1,-1,0)^T,\frac{1}{\sqrt{6}}(1,1,-2)^T,$$

于是化为第二标准型的正交变换为

$$R_2=\frac{1}{\sqrt{6}}\begin{bmatrix}\sqrt{2}&\sqrt{3}&1\\\sqrt{2}&-\sqrt{3}&1\\\sqrt{2}&0&-2\end{bmatrix},$$

$\boldsymbol{x}=R\boldsymbol{y}\rightarrow Q=\widetilde{\boldsymbol{y}}\widetilde{R}_2KR_2\boldsymbol{y}=\widetilde{\boldsymbol{y}}\,\mathrm{diag}[5,2,2]\boldsymbol{y}=5y_1^2+2y_2^2+2y_3^2$. 再作变换

$$\boldsymbol{y}=\mathrm{diag}\left[\frac{1}{\sqrt{5}},\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right]\boldsymbol{z}\rightarrow Q=z_1^2+z_2^2+z_3^2.$$

故使 Q 化为第一标准型的正交变换是

$$R_1=R_2\mathrm{diag}\left[\frac{1}{\sqrt{5}},\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right]=\frac{1}{2\sqrt{15}}\begin{bmatrix}2 & \sqrt{15} & \sqrt{5}\\ 2 & -\sqrt{15} & \sqrt{5}\\ 2 & 0 & -2\sqrt{5}\end{bmatrix}.$$

6.4 求二次型 $Q=2x_1x_2+2x_2x_3+x_3^2$ 的秩和惯性，并确定使 Q 化为标准型的坐标变换[1].

解 依次实施一些简单变换，可得

$$\begin{aligned}Q&=(x_3^2+2x_3x_2)+2x_1x_2\\&=(x_1'+x_2')^2-x_2'^2+2x_3'x_2'\\&=y_1^2-y_2^2+2y_2y_3\\&=z_1^2-z_2^2+z_3^2.\end{aligned}$$

于是，$\mathrm{rank}Q=r=3$，$\mathrm{In}Q=(2,1,0)$，$\mathrm{sig}Q=1$. 并且容易求得使 Q 化为标准型的变换为

$$P=\begin{bmatrix}1&0&0\\0&1&-1\\0&0&1\end{bmatrix}\begin{bmatrix}1&1&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}=\begin{bmatrix}0&1&1\\-1&1&0\\1&0&0\end{bmatrix}.$$

这样做避免了求解本征方程 $\lambda^3-\lambda^2-2\lambda+1=0$ 的困难.

这种方法称为 **Lagrange 法**，其一般步骤是：选出含某一变量的所有项，使之能配成平方，作变量变换后得

$$Q(x_1,\cdots,x_n)=\alpha_1y_1^2+Q_1(y_2,\cdots,y_n),$$

这里 $y_2,\cdots,y_n$ 实际为选剩变量的重新命名. 如此继续进行下去，即可得所需结果.

6.5 用迭代法求矩阵

$$A=\begin{bmatrix}5&-2&-4\\-2&2&2\\-4&2&5\end{bmatrix}$$

的(绝对值)最大本征值 λ_1 及对应本征矢 $\boldsymbol{x}^{(1)}$[2].

解 选 $\overset{(0)}{\boldsymbol{x}}=(1,0,0)^{\mathrm{T}}$，并依次作如下迭代得

① 参考：P. Lancaster, M. Tismenetsky, *Theory of Matrices*, Academic, New York, 1985, pp. 206—207.

② 参考：A. S. Deif, *Advanced Matrix Theory for Scientists and Engineers*, p. 93.

$$\overset{(1)}{x} = A\overset{(0)}{x} = (5, -2, -4)^T,$$
$$\overset{(2)}{x} = A\overset{(1)}{x} = (45, -22, -44)^T,$$
$$\overset{(3)}{x} = A\overset{(2)}{x} = (445, -222, -444)^T;$$

充分多次(这里 $m=3$)迭代后给出本征值和本征矢为

$$\lambda_1 \approx \frac{445}{45} \approx \frac{222}{22} \approx \frac{444}{44} \approx 10,$$

$$x^{(1)} = \frac{1}{3}(2, -1, -2)^T.$$

这方法称为**乘幂法**,其原理是:任一矢量均可用 A 的本征矢展开,$x = \sum_{i=1}^{n} c_i x^{(i)}$,经 r 次迭代后得

$$\overset{(r)}{x} = \sum_{i=1}^{n} c_i \lambda_i^r x^{(i)} = \lambda_1^r \left[c_1 x^{(1)} + \sum_{i=2}^{n} \left(\frac{\lambda_i}{\lambda_1}\right)^r c_i x^{(i)} \right] \approx \lambda_1^r c_1 x^{(1)},$$

$$\overset{(r)}{x} \approx \lambda_1 \overset{(r-1)}{x}, \quad \overset{(r)}{x} \approx \alpha_1 x^{(1)};$$

r 次迭代后的收敛速率与 $\left(\frac{\lambda_2}{\lambda_1}\right)^r \to 0$ 的速率相同(注意到本题中的$\frac{\lambda_2}{\lambda_1}=0.1$).

为求次大本征值 λ_2,构造新矩阵

$$A' = A - \lambda_1 x^{(1)} x^{(1)+}$$

A'与 A 除 $\lambda_1'=0$ 外,有相同本征值和本征矢,故可再次应用乘幂法求 λ_2 和 $x^{(2)}$.

6.6　用迭代法求矩阵

$$A = \begin{bmatrix} 6 & 3 & 1 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

的本征值和本征矢①.

解　设本征值 $|\lambda_1| \geqslant |\lambda_2| \geqslant |\lambda_3|$,相应本征矢为 $x^{(1)}, x^{(2)}, x^{(3)}$. 下面给出迭代的实用格式.

① 用**乘幂法**求最大本征值 λ_1 和本征矢 $x^{(1)}$.

迭代的基本循环是

$$\left(y = Ax,\ x = \frac{1}{k} y\right),$$

标度因子 k 是为了防止计算中所涉及之数变得太大或太小,可取为矢量 y 的范数 $\|y\|$

① 参考: W. Ledermann (Chief ed.), *Handbook of Applicable Mathematics*, vol. Ⅲ, R. F. Churchhouse (ed.) *Numerical Methods*, Wiley, New York, 1981, pp. 108—111.

$$k = \pm \|\boldsymbol{y}\|_2 = \pm (\boldsymbol{y}^\dagger \boldsymbol{y})^{1/2} \quad 或 \quad k = \pm \|\boldsymbol{y}\|_\infty = \pm \max_i |y_i|,$$

这保证 $\|\boldsymbol{x}\|_2 = 1$ 或 $\|\boldsymbol{x}\|_\infty = 1$；其前的±号取 $\boldsymbol{y}$ 的具有最大模数之分量 y_i 的正负号. 迭代充分多次后，$\boldsymbol{x}$ 和 k 分别收敛于 $\boldsymbol{x}^{(1)}$ 和 λ_1.

本题的前 6 次迭代结果如表 6.1.

表 6.1　乘幂法迭代计算表

r	0	1	2	3	4	5	6
		10	26.8	33.58	34.751	34.907	34.926
$\boldsymbol{y}$		6	14.9	18.47	19.092	19.171	19.181
		3	6.3	7.61	7.840	7.869	7.873
	1	3.3	4.25	4.413	4.433	4.436	4.436
$\boldsymbol{x}$	1	2	2.36	2.427	2.435	2.436	2.436
	1	1	1	1	1	1	1

计算给出 $\lambda_1 = 7.873$（至 10 位的准确结果为 7.872 983 346），$x^{(1)} = (4.436, 2.436, 1)^{\mathrm{T}}$. 标度因子取为 y_3. 由于 $\lambda_2 = 1$，$\lambda_2/\lambda_1 = 0.127$，近似的误差如 $(0.127)^r$ 那样趋于零，每多迭代一次近于提高一位准确度.

② 用**逆迭代法**求最小本征值 λ_3 和本征矢 $\boldsymbol{x}^{(3)}$.

对于非奇异矩阵 A，

$$A\boldsymbol{x} = \lambda\boldsymbol{x} \implies A^{-1}A\boldsymbol{x} = A^{-1}\lambda\boldsymbol{x} \implies \boldsymbol{x} = \lambda A^{-1}\boldsymbol{x} \implies A^{-1}\boldsymbol{x} = \lambda^{-1}\boldsymbol{x},$$

$$(A - hI)\boldsymbol{x} = (\lambda - h)\boldsymbol{x} \implies \boldsymbol{x} = (\lambda - h)(A - hI)^{-1}\boldsymbol{x} \implies (A - hI)^{-1}\boldsymbol{x} = (\lambda - h)^{-1}\boldsymbol{x};$$

即 A，A^{-1} 和 $(A - hI)^{-1}$ 具有共同本征矢，而其本征值则分别为 λ，λ^{-1} 和 $(\lambda - h)^{-1}$；这里 h 称为（本征值的）原点移位. 若 h 是最小本征值的近似，则可用乘幂法于 $(A - hI)^{-1}$，可加速收敛性.

逆迭代法的基本循环是

$$\left[\boldsymbol{y} = (A - hI)^{-1}\boldsymbol{x}, \boldsymbol{x} = \frac{1}{k}\boldsymbol{y}, h(新) = h(旧) + \frac{1}{k}\right],$$

即由已知 $\boldsymbol{x}$ 和 h（初始一般取为 0）解出 $\boldsymbol{y}$，再适当选取标度因子 k 如(1)，以及新的 h. 其收敛性 $\propto \left(\dfrac{\lambda_n - h}{\lambda_{n-1} - h}\right)^r$ 很快.

本题的迭代结果如表 6.2，其中开始两次迭代取 $h = 0$，以后利用 Rayleigh-Ritz 比 $\dfrac{(\boldsymbol{x}, A\boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})}$ 计算 h. 计算结果是：

$$h = \lambda_3 = 0.127\,017（准确值为 0.127\,016\,654），$$

$$\boldsymbol{x}^{(3)} = (-0.392\,27, 1, 0.696\,13)^{\mathrm{T}}.$$

表 6.2　逆迭代法计算表

r	0	1	2	3
		1	10	−23.371
$\boldsymbol{y}$		−3	−26	59.580
		3	19	−41.481
k		1	−26	59.580
	0	1	−0.381	−0.392 27
$\boldsymbol{x}$	0	−3	1.000	1.000 00
	1	3	−0.730	0.696 13
h	0	0	0.127	0.127 017

第 7 章　线性积分方程

7.1 积 分 方 程

7.1.1 定义和分类

(1) 定义　凡在积分号下含有未知函数的方程统称为**积分方程**；若此方程对未知函数是线性的，则称为**线性积分方程**.

(2) 分类　线性积分方程有以下几种形式：

$$\int_a^b \overset{\vee}{K}(x,\xi)\overset{?}{\varphi}(\xi)\mathrm{d}\xi = \overset{\vee}{f}(x),$$

$$\overset{?}{\varphi}(x) - \lambda\int_a^b \overset{\vee}{K}(x,\xi)\overset{?}{\varphi}(\xi)\mathrm{d}\xi = \overset{\vee}{f}(x),$$

$$\overset{\vee}{g}(x)\overset{?}{\varphi}(x) - \lambda\int_a^b \overset{\vee}{K}(x,\xi)\overset{?}{\varphi}(\xi)\mathrm{d}\xi = \overset{\vee}{f}(x);$$

它们分别称为**第一类、第二类、第三类 Fredholm 积分方程**；其中 $K(x,\xi)$，$f(x)$，$g(x)$为已知函数(打√号者)，在其变量的区间 $[a,b]$ 上连续，$\varphi(x)$为未知函数(打？号者)，λ 是参数. $K(x,\xi)$称为**积分方程的核**.

可以看出，第一类方程和第二类方程分别是第三类方程当 $g(x)=0$、$\lambda=-1$ 和 $g(x)=1$ 时的特殊情况.

另外，若当 $\xi>x$ 时令 $K(x,\xi)=0$ 则得可变上限的积分方程，称为 **Volterra 积分方程**. 例如，

$$\varphi(x) - \lambda\int_a^x K(x,\xi)\varphi(\xi)\mathrm{d}\xi = f(x),$$

相应称为**第二类 Volterra 积分方程**.

虽然如此，第一类和第二类 Fredholm 积分方程二者在本质上是有区别的，Volterra 积分方程与 Fredholm 积分方程也是本质上不同的.

7.1.2 对应无穷代数方程组

现在来考虑下列线性积分方程

$$\int_a^b A(u,v)X(v)\mathrm{d}v = C(u).$$

按照 Riemann 积分的概念，

$$\int_a^b f(v)\mathrm{d}v = \lim_{N\to\infty}\sum_{j=1}^{N} f(v_j)\delta,$$

其中 $\delta=(b-a)/N, v_j=a+j\delta(v_0=a, v_N=b)$；所以，前述积分方程可写成

$$\lim_{N\to\infty}\sum_{j=1}^{N} A(u,v_j)X(v_j)\delta = C(u).$$

由于此方程两边为区间 $a\leqslant u\leqslant b$ 上的连续函数，若令 $u_i=a+i\delta$，故有

$$\lim_{N\to\infty}\sum_{j=1}^{N} A(u_i,v_j)X(v_j)\delta = C(u_i) \quad (i=1,2,\cdots,N).$$

因此，前述线性积分方程相当于无穷($n=N\to\infty$)个未知量的线性代数方程组：

$$\sum_{j=1}^{n} a_{ij}x_j = c_i \quad (i=1,2,\cdots,n);$$

具有

$$a_{ij} = A(u_i,v_j)\delta, \quad x_j = X(v_j), \quad c_i = C(u_i).$$

显然，对于一般积分方程亦有类似结果.

下面将着重讨论第二类积分方程的解法.

7.2 第二类积分方程的 Fredholm 解[①]

Fredholm 解法的步骤是先求对应 N 元代数方程组的解，再求使 $N\to\infty$ 之极限.

7.2.1 对应代数方程组及其解法

与第二类 Fredholm 积分方程：

$$\varphi(x)-\lambda\int_a^b K(x,\xi)\varphi(\xi)\mathrm{d}\xi = f(x),$$

对应的线性代数方程为

$$\varphi(x_i)-\lambda\sum_{j=1}^{N} K(x_i,\xi_j)\varphi(\xi_j)\delta = f(x_i);$$

或者写成

$$\sum_{j=1}^{N} a_{ij}\varphi_j = f_i, \quad \text{或} \quad A\boldsymbol{\varphi} = \boldsymbol{f};$$

具有系数矩阵 $A=\{a_{ij}\}$，

---

① E. T. Whittaker, G. N. Watson, *Modern Analysis*, Cambridge University Press, Cambridge, 4th ed., 1940; §11.2.

$$a_{ij} = \delta_{ij} - \lambda K_{ij}\delta, \quad K_{ij} \equiv K(x_i, \xi_j);$$

以及

$$\varphi_i \equiv \varphi(x_i), \quad \varphi_j \equiv \varphi(\xi_j), \quad f_i \equiv f(x_i);$$

这里 $x_k = a + k\delta = \xi_k, \delta = (b-a)/N$.

(1) 行列式解法

上述代数方程组 $A\boldsymbol{\varphi} = \boldsymbol{f}$ 的解为 $\boldsymbol{\varphi} = A^{-1}\boldsymbol{f}$,即

$$\varphi_i = \sum_j (A^{-1})_{ij} f_j = \sum_j \frac{A_{ji}}{|A|} f_j,$$

其中 A_{ji} 是行列式 $|A|$ 中 a_{ji} 的代数余子式,

$$A_{ji} = \frac{\partial |A|}{\partial a_{ji}}.$$

所以,如能算出 $A_{ji}/|A|$,即可求得 φ_i.

(2) 行列式的级数展开

现在先把 $|A|$ 当做 N 阶行列式,按 λ 的幂作级数展开:

$$|A| = \begin{vmatrix} 1-\lambda K_{11}\delta & -\lambda K_{12}\delta & -\lambda K_{13}\delta & \cdots \\ -\lambda K_{21}\delta & 1-\lambda K_{22}\delta & -\lambda K_{23}\delta & \cdots \\ -\lambda K_{31}\delta & -\lambda K_{32}\delta & 1-\lambda K_{33}\delta & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} = 1 - \lambda \sum_j K_{jj}\delta + \cdots.$$

关于 λ^2 的系数:$N-2$ 个对角元用其 1,其余 2 个由 N 个中取 jk 对,有 $\binom{N}{2}$ 种取法;因为凡是行列式中对角元项不变时则行列式之值不变,故 λ^2 的系数为 $\frac{1}{2!}\sum_{j=1}^{N}\sum_{k=1}^{N}\begin{vmatrix} K_{jj} & K_{jk} \\ K_{kj} & K_{kk} \end{vmatrix}\delta^2$. 同样可得 λ^3 的系数. 所以得到 $|A|$ 的 λ 级数展开为

$$|A| = 1 - \lambda\sum_{j=1}^{N} K_{jj}\delta + \frac{\lambda^2}{2!}\sum_{j=1}^{N}\sum_{k=1}^{N}\begin{vmatrix} K_{jj} & K_{jk} \\ K_{kj} & K_{kk} \end{vmatrix}\delta^2 - \frac{\lambda^3}{3!}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\begin{vmatrix} K_{ii} & K_{ij} & K_{ik} \\ K_{ji} & K_{jj} & K_{jk} \\ K_{ki} & K_{kj} & K_{kk} \end{vmatrix}\delta^3 + \cdots.$$

7.2.2 Fredholm 行列式

现在来求上述行列式的极限,让 $N \to \infty, \delta \to 0$,则有

$$\sum_{j=1}^{N} K_{jj}\delta \to \int_a^b K(\xi_1, \xi_1)\,\mathrm{d}\xi_1,$$

$$\sum_{j=1}^{N}\sum_{k=1}^{N}\begin{vmatrix} K_{jj} & K_{jk} \\ K_{kj} & K_{kk} \end{vmatrix}\delta^2 \to \int_a^b\int_a^b \begin{vmatrix} K(\xi_1,\xi_1) & K(\xi_1,\xi_2) \\ K(\xi_2,\xi_1) & K(\xi_2,\xi_2) \end{vmatrix}\mathrm{d}\xi_1\,\mathrm{d}\xi_2,$$

等等. 于是, $|A|\to D(\lambda)$:

$$D(\lambda)=1-\lambda\int_a^b K(\xi_1,\xi_1)\mathrm{d}\xi_1+\frac{\lambda^2}{2!}\int_a^b\int_a^b\begin{vmatrix}K(\xi_1,\xi_1) & K(\xi_1,\xi_2)\\ K(\xi_2,\xi_1) & K(\xi_2,\xi_2)\end{vmatrix}\mathrm{d}\xi_1\mathrm{d}\xi_2$$

$$-\frac{\lambda^3}{3!}\int_a^b\int_a^b\int_a^b\begin{vmatrix}K(\xi_1,\xi_1) & K(\xi_1,\xi_2) & K(\xi_1,\xi_3)\\ K(\xi_2,\xi_1) & K(\xi_2,\xi_2) & K(\xi_2,\xi_3)\\ K(\xi_3,\xi_1) & K(\xi_3,\xi_2) & K(\xi_3,\xi_3)\end{vmatrix}\mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}\xi_3$$

$$+-\cdots \text{至无穷},$$

式中用"$+-$"表示各项正负交替. $D(\lambda)$称为 **Fredholm 行列式.**

现在来求代数余子式 $A_{ji}(j\neq i)$

$$A_{ji}=\frac{\partial|A|}{\partial a_{ji}}$$

$$=\lambda\delta\left[K_{ij}-\lambda\sum_k\begin{vmatrix}K_{ij} & K_{ik}\\ K_{kj} & K_{kk}\end{vmatrix}\delta+\frac{\lambda^2}{2!}\sum_k\sum_l\begin{vmatrix}K_{ij} & K_{ik} & K_{il}\\ K_{kj} & K_{kk} & K_{kl}\\ K_{lj} & K_{lk} & K_{ll}\end{vmatrix}\delta^2-\cdots\right];$$

当 $N\to\infty,\delta\to 0$ 时, $\dfrac{A_{ji}}{\delta}\to D(\xi_i,\xi_j|\lambda)$:

$$D(\xi_i,\xi_j|\lambda)\equiv\lambda\Big[K(\xi_i,\xi_j)-\lambda\int_a^b\begin{vmatrix}K(\xi_i,\xi_j) & K(\xi_i,\xi_1)\\ K(\xi_1,\xi_j) & K(\xi_1,\xi_1)\end{vmatrix}\mathrm{d}\xi_1$$

$$+\frac{\lambda^2}{2!}\int_a^b\int_a^b\begin{vmatrix}K(\xi_i,\xi_j) & K(\xi_i,\xi_1) & K(\xi_i,\xi_2)\\ K(\xi_1,\xi_j) & K(\xi_1,\xi_1) & K(\xi_1,\xi_2)\\ K(\xi_2,\xi_j) & K(\xi_2,\xi_1) & K(\xi_2,\xi_2)\end{vmatrix}\mathrm{d}\xi_1\mathrm{d}\xi_2$$

$$-+\cdots \text{至无穷}\Big].$$

$D(x,y|\lambda)$称为 **Fredholm 初余子式.**

再来求代数余子式 A_{ii}. 因为显然有$\dfrac{\partial|A|}{\partial a_{ii}}\to|A|$,所以有 $A_{ii}\to D(\lambda)$.

总结起来,在 $N\to\infty,\delta\to 0$ 极限下有

$$|A|\to D(\lambda),$$
$$A_{ii}\to D(\lambda),$$
$$A_{ji}\to D(x_i,y_j|\lambda)\delta(j\neq i).$$

7.2.3 Fredholm 解

有了以上结果,于是

$$\varphi_i=\sum_j\frac{A_{ji}}{|A|}f_j=\frac{A_{ii}}{|A|}f_i+\sum_{j\neq i}\frac{A_{ji}}{|A|}f_j,\quad \text{即}$$

$$\varphi_i = f_i + \sum_{j \neq i} \frac{D(x_i, y_j \mid \lambda)}{D(\lambda)} f_j \delta.$$

若 $D(\lambda) \neq 0$，相当于非奇异情况，则第二类积分方程之 **Fredholm** 解为

$$\varphi(x) = f(x) + \int_a^b \frac{D(x, y \mid \lambda)}{D(\lambda)} f(y) \mathrm{d}y.$$

(1) 倒易核

若令

$$-\frac{D(x, y \mid \lambda)}{D(\lambda)} = \lambda k(x, y \mid \lambda),$$

其中 $k(x, y|\lambda)$ 称为**倒易核**或**预解核**，则有

$$f(x) = \varphi(x) - \lambda \int_a^b K(x, \xi) \varphi(\xi) \mathrm{d}\xi,$$

$$\varphi(x) = f(x) - \lambda \int_a^b k(x, \xi \mid \lambda) f(\xi) \mathrm{d}\xi.$$

由此可见，若 $K(x, \xi)$，λ 及 $f(x)$ 为已知，前者是关于 $\varphi(x)$ 的积分方程，从给定的 K，λ 就可算出 $k(x, \xi|\lambda)$，则后者即为其解. 另一方面，若给定 $k(x, \xi)$，λ 及 $\varphi(x)$，后者可看作关于 $f(x)$ 的积分方程，同样可算出 $K(x, \xi)$，则前者给出其解. 因此，倒易核 $k(x, \xi)$ 之倒易核为核 $K(x, \xi)$；即 $K(x, \xi)$ 和 $k(x, \xi)$ 互为倒易核.

(2) Fredholm 基本关系

若将解式(后者)代入积分方程(前者)应得恒等式：

$$f(x) = \left[f(x) - \lambda \int_a^b k(x, \xi \mid \lambda) f(\xi) \mathrm{d}\xi \right] - \lambda \int_a^b K(x, \xi) \left[f(\xi) - \lambda \int_a^b k(\xi, y \mid \lambda) f(y) \mathrm{d}y \right] \mathrm{d}\xi.$$

化简后可简洁写成

$$\int kf + \int K \left[f - \lambda \int kf \right] \equiv 0 \quad \text{或} \quad \int [K + k - \lambda \int Kk] f \equiv 0,$$

对后一式中的重积分曾作积分次序和变量之交换. 我们知道，若对任意 f_i 有 $\sum X_i f_i \equiv 0$，则 $X_i = 0$. 现在对任意 $f(\xi)$ 有 $\int X(\xi) f(\xi) \mathrm{d}\xi \equiv 0$，所以 $X(\xi) = 0$. 故有倒易核之间的 **Fredholm 基本关系**：

$$\begin{aligned} K(x, \xi) + k(x, \xi \mid \lambda) &= \lambda \int_a^b K(x, y) k(y, \xi \mid \lambda) \mathrm{d}y \\ &= \lambda \int_a^b k(x, y \mid \lambda) K(y, \xi) \mathrm{d}y, \end{aligned}$$

后一等式是由于 K 和 k 互为倒易核.

(3) Fredholm 解

现在再将 **Fredholm** 解的结果列出如下：

积分方程

$$\varphi(x)-\lambda\int_a^b K(x,\xi)\varphi(\xi)\mathrm{d}\xi=f(x)$$

之解为

$$\varphi(x)=f(x)+\frac{1}{D(\lambda)}\int_a^b D(x,\xi\mid\lambda)f(\xi)\mathrm{d}\xi,$$

其中

$$D(\lambda)=1-\lambda\int_a^b K(\xi_1,\xi_1)\mathrm{d}\xi_1+\frac{\lambda^2}{2!}\int_a^b\int_a^b\begin{vmatrix}K(\xi_1,\xi_1) & K(\xi_1,\xi_2)\\ K(\xi_2,\xi_1) & K(\xi_2,\xi_2)\end{vmatrix}\mathrm{d}\xi_1\mathrm{d}\xi_2-+\cdots,$$

$$\begin{aligned}D(x,y\mid\lambda)=&\lambda K(x,y)-\lambda^2\int_a^b\begin{vmatrix}K(x,y) & K(x,\xi_1)\\ K(\xi_1,y) & K(\xi_1,\xi_1)\end{vmatrix}\mathrm{d}\xi_1\\&+\frac{\lambda^3}{2!}\int_a^b\int_a^b\begin{vmatrix}K(x,y) & K(x,\xi_1) & K(x,\xi_2)\\ K(\xi_1,y) & K(\xi_1,\xi_1) & K(\xi_1,\xi_2)\\ K(\xi_2,y) & K(\xi_2,\xi_1) & K(\xi_2,\xi_2)\end{vmatrix}\mathrm{d}\xi_1\mathrm{d}\xi_2\\&-+\cdots.\end{aligned}$$

(4) 可解条件

积分方程之可解条件是:

① $D(\lambda)\neq 0$,相当于非奇异情况.

② $D(\lambda)$,$D(x,y\mid\lambda)$ 收敛. 从有限变到无穷,出现收敛性问题,需要证明: $K(x,y)$在什么条件下,能使 $D(\lambda)$,$D(x,y|\lambda)$收敛? 如果收敛才是解,最方便的方法就是代进去检验一下. 具体说,若在区间 $a\leqslant x,y\leqslant b$ 上 $K(x,y)$为连续实变函数(Fredholm 方程)或者在区间 $a\leqslant y\leqslant x\leqslant b$ 上 $K(x,y)$为连续实变函数,而当 $y>x$ 时 $K(x,y)=0$(Volterra 方程),则方程可解. 关键是 $K(x,y)$一定要有限.

7.2.4 例子

下面举一个具体例子:

$$\varphi(x)-\lambda\int_0^1(x\xi+\xi^2)\varphi(\xi)\mathrm{d}\xi=x;$$

这里 $K(x,y)=xy+y^2=y(x+y)$,$f(x)=x$. 于是,

$$K(\xi_1,\xi_1)=2\xi_1^2,$$

$$\begin{vmatrix}K(\xi_1,\xi_1) & K(\xi_1,\xi_2)\\ K(\xi_2,\xi_1) & K(\xi_2,\xi_2)\end{vmatrix}=\xi_1\xi_2\begin{vmatrix}2\xi_1 & \xi_1+\xi_2\\ \xi_2+\xi_1 & 2\xi_2\end{vmatrix}=-\xi_1\xi_2(\xi_1-\xi_2)^2,$$

当 $n\geqslant 3$ 时,各行列式均为 0,因为

$$\begin{vmatrix} K(\xi_1,\xi_1) & K(\xi_1,\xi_2) & K(\xi_1,\xi_3) & \cdots & K(\xi_1,\xi_n) \\ K(\xi_2,\xi_1) & K(\xi_2,\xi_2) & K(\xi_2,\xi_3) & \cdots & K(\xi_2,\xi_n) \\ K(\xi_3,\xi_1) & K(\xi_3,\xi_2) & K(\xi_3,\xi_3) & \cdots & K(\xi_3,\xi_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ K(\xi_n,\xi_1) & K(\xi_n,\xi_2) & K(\xi_n,\xi_3) & \cdots & K(\xi_n,\xi_n) \end{vmatrix}$$

$$= \xi_1\xi_2\xi_3\cdots\xi_n \begin{vmatrix} \xi_1+\xi_1 & \xi_1+\xi_2 & \xi_1+\xi_3 & \cdots & \xi_1+\xi_n \\ \xi_2+\xi_1 & \xi_2+\xi_2 & \xi_2+\xi_3 & \cdots & \xi_2+\xi_n \\ \xi_3+\xi_1 & \xi_3+\xi_2 & \xi_3+\xi_3 & \cdots & \xi_3+\xi_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \xi_n+\xi_1 & \xi_n+\xi_2 & \xi_n+\xi_3 & \cdots & \xi_n+\xi_n \end{vmatrix}$$

(后面各列减去第一列后得下式)

$$= \xi_1\xi_2\xi_3\cdots\xi_n \begin{vmatrix} \xi_1+\xi_1 & \xi_2-\xi_1 & \xi_3-\xi_1 & \cdots & \xi_n-\xi_1 \\ \xi_2+\xi_1 & \xi_2-\xi_1 & \xi_3-\xi_1 & \cdots & \xi_n-\xi_1 \\ \xi_3+\xi_1 & \xi_2-\xi_1 & \xi_3-\xi_1 & \cdots & \xi_n-\xi_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \xi_n+\xi_1 & \xi_2-\xi_1 & \xi_3-\xi_1 & \cdots & \xi_n-\xi_1 \end{vmatrix}$$

(后面各列提出因子后全同得下面结果)

$$= 0.$$

因此,可得

$$D(\lambda) = 1-\lambda\int_0^1 K(\xi_1,\xi_1)\mathrm{d}\xi_1 + \frac{\lambda^2}{2!}\int_0^1\int_0^1 \begin{vmatrix} K(\xi_1,\xi_1) & K(\xi_1,\xi_2) \\ K(\xi_2,\xi_1) & K(\xi_2,\xi_2) \end{vmatrix} \mathrm{d}\xi_1\mathrm{d}\xi_2$$

$$= 1-\lambda\int_0^1 2\xi_1^2\mathrm{d}\xi_1 - \frac{\lambda^2}{2}\int_0^1\int_0^1 \xi_1\xi_2(\xi_1-\xi_2)^2\mathrm{d}\xi_1\mathrm{d}\xi_2$$

$$= 1-\frac{2}{3}\lambda-\frac{1}{72}\lambda^2.$$

类似地,可得

$$D(x,y\,|\,\lambda) = \lambda K(x,y) - \lambda^2\int_0^1 \begin{vmatrix} K(x,y) & K(x,\xi_1) \\ K(\xi_1,y) & K(\xi_1,\xi_1) \end{vmatrix} \mathrm{d}\xi_1$$

$$= \lambda y(x+y) - \lambda^2\int_0^1 [y(x+y)\cdot 2\xi_1^2 - y\xi_1(\xi_1+x)(\xi_1+y)]\mathrm{d}\xi_1$$

$$= \lambda(xy+y^2) + \lambda^2\left[\left(\frac{1}{2}x-\frac{1}{3}\right)y^2 - \left(\frac{1}{3}x-\frac{1}{4}\right)y\right].$$

最后求得方程之解为

$$\varphi(x) = x + \frac{1}{D(\lambda)}\int_0^1 D(x,y\,|\,\lambda)y\mathrm{d}y = \frac{(72-24\lambda)x+18\lambda}{72-48\lambda-\lambda^2}.$$

7.3 第二类积分方程的 Liouville 迭代解

将积分方程写成

$$\varphi(x) = f(x) + \lambda\int_a^b K(x,\xi)\varphi(\xi)\mathrm{d}\xi,$$

用迭代法求解可得

零级近似： $\varphi(x)=f(x)\equiv\varphi_0(x)$，

一级近似： $\varphi(x) = f(x) + \lambda\int_a^b K(x,\xi)\varphi_0(\xi)\mathrm{d}\xi \equiv \varphi_1(x)$，

二级近似： $\varphi(x) = f(x) + \lambda\int_a^b K(x,\xi)\varphi_1(\xi)\mathrm{d}\xi \equiv \varphi_2(x)$，

$$\vdots$$

$n+1$ 级近似： $\varphi(x) = f(x) + \lambda\int_a^b K(x,\xi)\varphi_n(\xi)\mathrm{d}\xi \equiv \varphi_{n+1}(x)$.

或者，将 $\varphi_2(x)$展开可得

$$\begin{aligned}\varphi_2(x) &= f(x) + \lambda\int_a^b K(x,\xi)\left[f(\xi) + \lambda\int_a^b K(\xi,y)f(y)\mathrm{d}y\right]\mathrm{d}\xi \\ &= f(x) + \lambda\int_a^b K(x,\xi)f(\xi)\mathrm{d}\xi + \lambda^2\int_a^b K^{(2)}(x,y)f(y)\mathrm{d}y,\end{aligned}$$

其中曾交换积分次序并令

$$K^{(2)}(x,y) = \int_a^b K(x,\xi)K(\xi,y)\mathrm{d}\xi.$$

类似地可令

$$K^{(3)}(x,y) = \int_a^b K(x,\xi)K^{(2)}(\xi,y)\mathrm{d}\xi,$$

$$\vdots$$

$$K^{(n+1)}(x,y) = \int_a^b K(x,\xi)K^{(n)}(\xi,y)\mathrm{d}\xi;$$

于是，可得 **Liouville** 迭代解为

$$\begin{aligned}\varphi(x) &= f(x) + \int_a^b [\lambda K(x,\xi) + \lambda^2 K^{(2)}(x,\xi) + \lambda^3 K^{(3)}(x,\xi) + \cdots]f(\xi)\mathrm{d}\xi \\ &= f(x) + \int_a^b \Big[\sum_n \lambda^n K^{(n)}(x,\xi)\Big]f(\xi)\mathrm{d}\xi.\end{aligned}$$

用这方法不容易得到闭型精确解. 一般情况下，被积函数所包括的级数项有无限个. 假定此级数收敛，则会有

$$\frac{D(x,y\,|\,\lambda)}{D(\lambda)} = \sum_n \lambda^n K^{(n)}(x,y).$$

7.4 齐次积分方程

7.4.1 有非平凡解的条件

对于积分方程

$$\varphi(x)-\lambda\int_a^b K(x,\xi)\varphi(\xi)\mathrm{d}\xi=f(x),$$

其对应代数方程组为

$$\begin{bmatrix} 1-\lambda K(x_1,x_1)\delta & -\lambda K(x_1,x_2)\delta & \cdots & -\lambda K(x_1,x_N)\delta \\ -\lambda K(x_2,x_1)\delta & 1-\lambda k(x_2,x_2)\delta & \cdots & -\lambda K(x_2,x_N)\delta \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda K(x_N,x_1)\delta & -\lambda K(x_N,x_2)\delta & \cdots & 1-\lambda K(x_N,x_N)\delta \end{bmatrix}\begin{bmatrix}\varphi(x_1)\\ \varphi(x_2)\\ \vdots \\ \varphi(x_N)\end{bmatrix}=\begin{bmatrix}f(x_1)\\ f(x_2)\\ \vdots \\ f(x_N)\end{bmatrix},$$

即

$$A\boldsymbol{\varphi}=\boldsymbol{f}.$$

如果$|A|=D(\lambda)=0$,则为奇异情况;在此情况,$[\varphi(x_1),\cdots,\varphi(x_N)]^{\mathrm{T}}$ 被映射到其子空间中.因此,若$[f(x_1),\cdots,f(x_N)]^{\mathrm{T}}$ 是在 $\boldsymbol{\varphi}$ 的该子空间中,则方程有解,否则无解.

因为 $\boldsymbol{f}(x)=\boldsymbol{0}$ 总是在 $\boldsymbol{\varphi}$ 的子空间中,故**齐次积分方程**

$$\varphi(x)-\lambda\int_a^b K(x,\xi)\varphi(\xi)\mathrm{d}\xi=0$$

有非平凡解 $\varphi(x)\neq 0$ 的充要条件是 $D(\lambda)=0$.

7.4.2 看作本征值问题

将齐次积分方程写成

$$\int_a^b K(x,\xi)\varphi(\xi)\mathrm{d}\xi=\frac{1}{\lambda}\varphi(x),$$

其对应代数方程组为

$$\begin{bmatrix} K_{11}\delta & K_{12}\delta & \cdots & K_{1N}\delta \\ K_{21}\delta & K_{22}\delta & \cdots & K_{2N}\delta \\ \vdots & \vdots & \ddots & \vdots \\ K_{N1}\delta & K_{N2}\delta & \cdots & K_{NN}\delta \end{bmatrix}\begin{bmatrix}\varphi_1\\ \varphi_2\\ \vdots \\ \varphi_N\end{bmatrix}=\frac{1}{\lambda}\begin{bmatrix}\varphi_1\\ \varphi_2\\ \vdots \\ \varphi_N\end{bmatrix},$$

其中 $K_{ij}=K(x_i,x_j),\varphi_j=\varphi(x_j),\delta=(b-a)/N,x_k=a+k\delta$;即

$$\sum(K_{ij}\delta)\varphi_j=\frac{1}{\lambda}\varphi_i \quad \text{或} \quad \mathscr{K}\boldsymbol{\varphi}=\frac{1}{\lambda}\boldsymbol{\varphi},$$

后式中矩阵 $\mathscr{K}=\{K_{ij}\delta\}$.这就与第6章中讨论过的本征值问题一样.但应注意,积分

方程的本征值常用$\frac{1}{\lambda}$表示，如上式所示. 这与矩阵的本征值用 λ 表示，写为 $K\varphi=\lambda\varphi$ 的习惯不同.

现在的问题是要求出 λ 的这样一些值 λ_n，它们使上述齐次积分方程具有非平凡解；这些 λ_n 值以及相应的解 $\varphi^{(n)}(x)$ 分别称为核 $K(x,y)$ 的本征值和本征函数.

7.4.3 对称核与 Schmidt 定理

若积分方程的核是对称的，$K(x,y)=K(y,x)$，则上述 $\mathscr{K}=\{K_{ij}\delta\}$ 是对称矩阵. 根据本征值问题的理论，有本征值和本征矢存在. 因此，在极限($\delta\to 0$)情况下，上述齐次积分方程在例如 $\lambda=\lambda_n$ 时有解 $\varphi^{(n)}(x)$，即

$$\int_a^b K(x,\xi)\varphi^{(n)}(\xi)\,\mathrm{d}\xi=\frac{1}{\lambda_n}\varphi^{(n)}(x),$$

这里 λ_n 称为核 $K(x,y)$ 的本征值，而相应的 $\varphi^{(n)}(x)$ 称为其本征函数.

这样一来，也就证明了下述 **Schmidt 定理**：若积分方程的核是对称的，则前述 Fredholm 行列式为零的方程 $D(\lambda)=0$ 至少有一个根.

7.4.4 本征函数的正交归一化

设 $\lambda_1,\lambda_2,\cdots,\lambda_m,\cdots,\lambda_n,\cdots$ 为 $D(\lambda)=0$ 的根，$\varphi^{(1)}(x),\varphi^{(2)}(x),\cdots,\varphi^{(m)}(x),\cdots,\varphi^{(n)}(x),\cdots$ 是与这些本征值相对应的本征函数，则它们具有下列性质.

(i) 与不同本征值相对应的本征函数是正交的.

证 若 $\lambda_m\neq\lambda_n$，于是

$$\int_a^b K(x,\xi)\varphi^{(m)}(\xi)\,\mathrm{d}\xi=\frac{1}{\lambda_m}\varphi^{(m)}(x),$$

$$\int_a^b K(x,\xi)\varphi^{(n)}(\xi)\,\mathrm{d}\xi=\frac{1}{\lambda_n}\varphi^{(n)}(x).$$

对上列二式分别施加运算 $\int_a^b\cdots\varphi^{(n)}(x)\,\mathrm{d}x$ 和 $\int_a^b\cdots\varphi^{(m)}(x)\,\mathrm{d}x$ 于其两边，得到

$$\int_a^b\int_a^b\varphi^{(n)}(x)K(x,\xi)\varphi^{(m)}(\xi)\,\mathrm{d}\xi\mathrm{d}x=\frac{1}{\lambda_m}\int_a^b\varphi^{(n)}(x)\varphi^{(m)}(x)\,\mathrm{d}x,$$

$$\int_a^b\int_a^b\varphi^{(m)}(x)K(x,\xi)\varphi^{(n)}(\xi)\,\mathrm{d}\xi\mathrm{d}x=\frac{1}{\lambda_n}\int_a^b\varphi^{(m)}(x)\varphi^{(n)}(x)\,\mathrm{d}x.$$

首先注意到上面二式左边是相等的，因为只要在其一中交换积分变量 $x\to\xi$ 并计及积分方程核的对称 $K(\xi,x)=K(x,\xi)$ 就立即可以看出这点. 于是，上面二式相减后得到

$$\left(\frac{1}{\lambda_m}-\frac{1}{\lambda_n}\right)\int_a^b\varphi^{(m)}(x)\varphi^{(n)}(x)\,\mathrm{d}x=0;$$

由于设定 $\lambda_m \neq \lambda_n$，所以

$$\int_a^b \varphi^{(m)}(x)\varphi^{(n)}(x)\mathrm{d}x = 0.$$

这就是二函数 $\varphi^{(m)}(x)$，$\varphi^{(n)}(x)$ 相互正交的定义，它是第6章关于本征值问题的讨论中与不同本征值相对应的本征矢相互正交这一性质的直接推广，即 $\boldsymbol{\varphi}^{(m)}\cdot\boldsymbol{\varphi}^{(n)}=0$ 或

$$[\varphi^{(m)}(x_1)\varphi^{(n)}(x_1)+\varphi^{(m)}(x_2)\varphi^{(n)}(x_2)+\cdots+\varphi^{(m)}(x_N)\varphi^{(n)}(x_N)]\delta = 0$$

的极限性质.

(ii) 本征函数可以选择为归一的，即

$$\int_a^b \varphi^{(n)}(x)\varphi^{(n)}(x)\mathrm{d}x = 1.$$

这与本征值问题中选择本征矢为单位矢 $\boldsymbol{\varphi}^{(n)}\cdot\boldsymbol{\varphi}^{(n)}=1$ 相当.

换句话说，当积分方程的核为对称函数 $K(x,y)=K(y,x)$ 时，齐次积分方程

$$\varphi(x)-\lambda\int_a^b K(x,\xi)\varphi(\xi)\mathrm{d}\xi = 0$$

可当做无穷维空间中的主轴变换来看待.对特定的 λ 值，它有固定的方向与长度，分别称为本征值和本征矢(本征函数)，以及其他一些类似性质.

以上的正交与归一性质可结合起来写成

$$\int_a^b \varphi^{(m)}(x)\varphi^{(n)}(x)\mathrm{d}x = \delta_{mn},$$

其中 δ_{mn} 是 Kronecker 符号.满足这种性质的函数系列称为**正交归一本征函数系**.

7.4.5 求本征值和本征矢的 Aitken 方法

设有齐次线性方程组

$$\sum a_{ij}x_j = \lambda x_i \quad \text{或} \quad A\boldsymbol{x} = \lambda\boldsymbol{x},$$

其中方阵 $A=\{a_{ij}\}$.令其本征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，并且 $|\lambda_1|>|\lambda_2|>\cdots>|\lambda_n|$，相应正交归一本征矢为 $\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\cdots,\boldsymbol{x}^{(n)}$.因为任意矢量 $\boldsymbol{x}^{\mathrm{arb}}$ 均可用本征矢作线性展开：

$$\boldsymbol{x}^{\mathrm{arb}} = \sum_l \alpha_l \boldsymbol{x}^{(l)},$$

当用 A^k 作用于 $\boldsymbol{x}^{\mathrm{arb}}$ 时，若 k 充分大，则有

$$A^k\boldsymbol{x}^{\mathrm{arb}} = \sum_l \alpha_l A^k \boldsymbol{x}^{(l)} = \sum_l \alpha_l \lambda_l^k \boldsymbol{x}^{(l)} \approx \alpha_1\lambda_1^k\boldsymbol{x}^{(1)},$$

即其结果近似为(或正比于)最大本征矢；由此可求出 λ_1 和 $\boldsymbol{x}^{(1)}$.根据这一事实，A. C. Aitken 提出求本征值和本征矢的一种实用方法，其具体作法如下：

首先作出方阵 A 的下列平方幂序列

$$AA = A^2, A^2A^2 = A^4, A^4A^4 = A^8, \cdots, A^{2^k}A^{2^k} = A^{2^{k+1}}, \cdots;$$

将其作用于一任意矢量 $\boldsymbol{x}^{\text{arb}}$，则得试用矢量 $\boldsymbol{x}^{\text{trial}}$，例如

$$A^{2^k}\boldsymbol{x}^{\text{arb}}=\boldsymbol{x}^{\text{trial}}\propto\boldsymbol{x}^{(1)},$$

由此代回可求出最大本征值 λ_1 及相应本征矢 $\boldsymbol{x}^{(1)}$，如果还不成的话，可再运算一次

$$A^{2^k}\boldsymbol{x}^{\text{trial}}=A^{2^{k+1}}\boldsymbol{x}^{\text{arb}}\propto\boldsymbol{x}^{(1)}.$$

或者，比较 $A^{2^{k+1}}\boldsymbol{x}^{\text{arb}}$ 与 $A^{2^k}\boldsymbol{x}^{\text{arb}}$，看其是否近似成正比；如果是的话，则可通过

$$[A^{2^{k+1}}\boldsymbol{x}^{\text{arb}}/A^{2^k}\boldsymbol{x}^{\text{arb}}]^{2^{-k}}$$

近似求出 λ_1，而 $\boldsymbol{x}^{\text{trial}}$的归一化矢量即为 $\boldsymbol{x}^{(1)}$. 否则的话，需再试算下去.

这样求出最大本征值 λ_1 及相应本征矢 $\boldsymbol{x}^{(1)}$ 后，如何求其余的本征值和本征矢呢？这里要用到对称矩阵的下列展开形式.

设有对称矩阵 $A=\{a_{ij}\}$，其本征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，并且 $|\lambda_1|>|\lambda_2|>\cdots>|\lambda_n|$，其相应正交归一本征矢为 $\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\cdots,\boldsymbol{x}^{(n)}$. 前述齐次线性方程组还可用张量方程的形式表示为

$$\mathsf{A}\cdot\boldsymbol{x}=\lambda\boldsymbol{x}.$$

其中张量 A 的分量即为矩阵元 a_{ij}，左边的“$\cdot$”表示“缩乘”. 于是，张量 A 可用本征值和本征矢展开为

$$\mathsf{A}=\sum_l\lambda_l\boldsymbol{x}^{(l)}\boldsymbol{x}^{(l)}\quad \text{即}\quad a_{ij}=\sum_l\lambda_l x_i^{(l)}x_j^{(l)}.$$

验证如下：

$$[\mathsf{A}\cdot\boldsymbol{x}^{(m)}]_i=\sum_j a_{ij}x_j^{(m)}=\sum_j\sum_l\lambda_l x_i^{(l)}x_j^{(l)}x_j^{(m)}$$
$$=\sum_l\lambda_l x_i^{(l)}\delta_{lm}=\lambda_m x_i^{(m)},$$

即

$$\mathsf{A}\cdot\boldsymbol{x}^{(m)}=\lambda_m\boldsymbol{x}^{(m)}.$$

若已求出最大本征值 λ_1 及相应本征矢 $\boldsymbol{x}^{(1)}$，考虑下列系数张量 A'：

$$\mathsf{A}'=\mathsf{A}-\lambda_1\boldsymbol{x}^{(1)}\boldsymbol{x}^{(1)}=\sum_{l=2}^{n}\lambda_l\boldsymbol{x}^{(l)}\boldsymbol{x}^{(l)},$$

即

$$a_{ij}'=a_{ij}-\lambda_1x_i^{(1)}x_j^{(1)}=\sum_{l=2}^{n}\lambda_l x_i^{(l)}x_j^{(l)};$$

显然，A'或 $A'=\{a_{ij}'\}$的本征值为 $\lambda_2,\cdots,\lambda_n$，其相应本征矢为 $\boldsymbol{x}^{(2)},\cdots,\boldsymbol{x}^{(n)}$. 于是，可仿照前面的作法求出次大本征值 λ_2 及相应本征矢 $\boldsymbol{x}^{(2)}$.

其余的本征值及本征矢可继续照此办理而全部求出.

以上方法因为求 A^{2^k} 等可在计算机上直接运算求得，所以很实用. 这种方法称为 **Aitken 方法**；由于采用对系数矩阵的逐次平方，又称**逐次平方法**.

7.4.6　齐次积分方程的本征函数系

对于具有对称核的齐次积分方程

$$\int_a^b K(x,\xi)\varphi(\xi)\mathrm{d}\xi = \frac{1}{\lambda}\varphi(x),$$

同样可用 Aitken 方法来求其本征值及正交归一本征函数系. 只需注意到下列对应关系(用"≙"符号表示"相当于"):

$$Ax = \lambda x \triangleq \int_a^b K(x,\xi)\varphi(\xi)\mathrm{d}\xi = \frac{1}{\lambda}\varphi(x),$$

$$\lambda \triangleq \frac{1}{\lambda},$$

$$x \triangleq \varphi(x),$$

$$A \triangleq K(x,y),$$

$$A^2 \triangleq K^2(x,y) = \int_a^b K(x,\xi)K(\xi,y)\mathrm{d}\xi,$$

$$A^4 \triangleq K^4(x,y) = \int_a^b K^2(x,\xi)K^2(\xi,y)\mathrm{d}\xi,$$

$$\vdots$$

$$A^{2^{k+1}} \triangleq K^{2^{k+1}}(x,y) = \int_a^b K^{2^k}(x,\xi)K^{2^k}(\xi,y)\mathrm{d}\xi,$$

等等;以及下列对应的展开形式

$$\mathsf{A} = \sum_l \lambda_l x^{(l)} x^{(l)} \triangleq K(x,y) = \sum \frac{1}{\lambda_l}\varphi^{(l)}(x)\varphi^{(l)}(y),$$

后者称为 **Hilbert-Schmidt 展开**. 这样就可以类似上一小节那样逐次求出 λ_1, $\varphi^{(1)}(x)$;λ_2,$\varphi^{(2)}(x)$;…. 当然,这里最先求出的实际上是最小本征值 λ_1. 然后采用

$$K'(x,y) = K(x,y) - \frac{1}{\lambda_1}\varphi^{(1)}(x)\varphi^{(1)}(y),$$

又可求出次最小 λ_2 及 $\varphi^{(2)}(x)$. 依此类推即可求出全体本征值及其本征函数系.

7.5　第二类积分方程的 Hilbert-Schmidt 解法

对具有对称核的第二类积分方程

$$\varphi(x) = f(x) + \lambda\int_a^b K(x,\xi)\varphi(\xi)\mathrm{d}\xi,$$

可采用 **Hilbert-Schmidt 解法**,具体步骤如下:

首先,利用上节的 Aitken 方法求出相应齐次积分方程的本征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$,

…及其正交归一本征函数系 $\varphi^{(1)}(x),\varphi^{(2)}(x),\cdots,\varphi^{(n)}(x),\cdots$.

其次，假定 $f(x)$ 可用 $\varphi^{(n)}(x)$ 展开为

$$f(x)=\sum_n c_n\varphi^{(n)}(x),$$

可确定出展开系数 c_n 为

$$c_n=\int_a^b\varphi^{(n)}(x)f(x)\mathrm{d}x.$$

最后，假定 $\varphi(x)$ 可展开为

$$\varphi(x)=\sum_n z_n\varphi^{(n)}(x),$$

代回积分方程，并考虑到 Hilbert-Schmidt 展开

$$K(x,\xi)=\sum_m\frac{1}{\lambda_m}\varphi^{(m)}(x)\varphi^{(m)}(\xi),$$

得到确定展开系数 z_l 的下列方程

$$\begin{aligned}\sum_n z_n\varphi^{(n)}(x)&=\sum_n c_n\varphi^{(n)}(x)+\lambda\int_a^b\Big(\sum_m\frac{1}{\lambda_m}\varphi^{(m)}(x)\varphi^{(m)}(\xi)\Big)\Big(\sum_n z_n\varphi^{(n)}(\xi)\Big)\mathrm{d}\xi\\&=\sum_n c_n\varphi^{(n)}(x)+\lambda\sum_n\frac{1}{\lambda_n}z_n\varphi^{(n)}(x);\end{aligned}$$

于是，用比较系数法或两边同乘以 $\varphi^{(l)}(x)$ 再积分容易得到

$$z_l=c_l+\frac{\lambda}{\lambda_l}z_l,$$

即

$$z_l=\Big(1-\frac{\lambda}{\lambda_l}\Big)^{-1}c_l=\Big(\frac{\lambda_l}{\lambda_l-\lambda}\Big)c_l.$$

因为 λ,λ_l,c_l 均为已知，故得到 $\varphi(x)$ 用 $\varphi^{(l)}(x)$ 展开的级数解，

$$\varphi(x)=\sum_l\Big(\frac{\lambda_l}{\lambda_l-\lambda}\Big)c_l\varphi^{(l)}(x).$$

这里有几点需要注意. 首先是要求 $D(\lambda)\neq0$，即 λ 不是本征值，这是 7.2.3 节提过的积分方程的可解条件. 其次是假定 $f(x)$ 可用 $\varphi^{(n)}(x)$ 作展开. 还有重要的是积分方程核 K 要满足某些条件，使前面的 Hilbert-Schmidt 展开成立. 总之，求出解后把它代回原积分方程进行检验，对就成了.

最后应该注意，以上所介绍的均为一般方法. 另外，对于积分方程，变量有多少没有分别，做法是一样的；不像微分方程有“常”和“偏”之分.

习　题

7.1　求下列齐次积分方程

$$\varphi(x)=\lambda\int_0^1(xy+y^2)\varphi(y)\mathrm{d}y$$

的本征值 λ_n 和本征函数 $\varphi^{(n)}(x)\{n=1,2\}$.

解 (a) 这里 $K(x,y)=xy+y^2$，由 7.2.4 节例子已求出 Fredholm 行列式 $D(\lambda)$ 为

$$D(\lambda)=1-\frac{2}{3}\lambda-\frac{1}{72}\lambda^2.$$

求 $D(\lambda)=0$ 的根可得本征值，由

$$1-\frac{2}{3}\lambda-\frac{1}{72}\lambda^2=0 \quad \text{即} \quad \lambda^2+48\lambda-72=0$$

求得

$$\lambda_1=-24+18\sqrt{2}, \quad \lambda_2=-24-18\sqrt{2}.$$

这样只能求出本征值，本征函数还是不能由此得出，所以应该换用别法.

(b) 试验法. 根据对核 $K(x,y)$ 的观察检查，取 $\varphi(x)=\alpha x+\beta$ 这一试用形式，代入积分方程进行求解和验证. 于是，

$$\begin{aligned}\alpha x+\beta &= \lambda\int_0^1(xy+y^2)(\alpha y+\beta)\mathrm{d}y\\ &=\lambda\int_0^1[(\alpha y^2+\beta y)x+(\alpha y^3+\beta y^2)]\mathrm{d}y\\ &=\lambda\left(\frac{1}{3}\alpha+\frac{1}{2}\beta\right)x+\lambda\left(\frac{1}{4}\alpha+\frac{1}{3}\beta\right),\end{aligned}$$

比较系数得到确定 α,β 的方程组为

$$\begin{cases}\alpha=\lambda\left(\dfrac{\alpha}{3}+\dfrac{\beta}{2}\right),\\ \beta=\lambda\left(\dfrac{\alpha}{4}+\dfrac{\beta}{3}\right);\end{cases} \quad \text{或} \quad \begin{cases}\left(1-\dfrac{\lambda}{3}\right)\alpha-\dfrac{\lambda}{2}\beta=0,\\ -\dfrac{\lambda}{4}\alpha+\left(1-\dfrac{\lambda}{3}\right)\beta=0;\end{cases}$$

有非平凡解的条件是

$$D(\lambda)=\begin{vmatrix}1-\dfrac{\lambda}{3} & -\dfrac{\lambda}{2}\\ -\dfrac{\lambda}{4} & 1-\dfrac{\lambda}{3}\end{vmatrix}=1-\frac{2}{3}\lambda-\frac{1}{72}\lambda^2=0.$$

由此求出本征值 λ_1,λ_2 为

$$\lambda_1=-24+18\sqrt{2}, \quad \lambda_2=-24-18\sqrt{2};$$

分别代入确定 α,β 的方程组，令 $\beta_1=-1,\beta_2=1$，则得到相应未归一化本征函数 $\varphi^{(1)}(x),\varphi^{(2)}(x)$ 为

$$\varphi^{(1)}(x)=\sqrt{2}x+1, \quad \varphi^{(2)}(x)=\sqrt{2}x-1.$$

7.2 求下列积分方程

$$\varphi(x)=x+\lambda\int_0^1(xy+y^2)\varphi(y)\mathrm{d}y$$

的解，其中 $\lambda\neq$本征值.

解　7.2.4 节中已将上述方程作为例子，用 Fredholm 方法解过，现在用 Hilbert-Schmidt 方法和试验法进行求解.

(a) Hilbert-Schmidt 方法. 由上题已求出齐次积分方程的本征值和本征函数分别为

$$\lambda_1=-24+18\sqrt{2},\quad \lambda_2=-24-18\sqrt{2};$$
$$\varphi^{(1)}(x)=\sqrt{2}x+1,\quad \varphi^{(2)}(x)=\sqrt{2}x-1.$$

容易求得

$$f(x)=x=\frac{\sqrt{2}}{4}[\varphi^{(1)}(x)+\varphi^{(2)}(x)],\text{即 } c_1=c_2=\frac{\sqrt{2}}{4};$$

因而 $\varphi(x)$可展开为

$$\varphi(x)=z_1\varphi^{(1)}(x)+z_2\varphi^{(2)}(x),$$
$$z_l=\left(\frac{\lambda_l}{\lambda_l-\lambda}\right)c_l\quad (l=1,2);$$

以 λ_l 和 c_l 的值代入后求得

$$z_1=\frac{6\sqrt{2}-9}{\lambda+24-18\sqrt{2}}=\frac{(6\sqrt{2}-9)\lambda-18\sqrt{2}}{\lambda^2+48\lambda-72},$$

$$z_2=\frac{6\sqrt{2}+9}{\lambda+24+18\sqrt{2}}=\frac{(6\sqrt{2}+9)\lambda-18\sqrt{2}}{\lambda^2+48\lambda-72};$$

于是，求得积分方程之解为

$$\varphi(x)=(z_1+z_2)\sqrt{2}x+(z_1-z_2)=\frac{(24\lambda-72)x-18\lambda}{\lambda^2+48\lambda-72}.$$

(b) 试验法. 取 $\varphi(x)=\alpha x+\beta$，则有

$$\alpha x+\beta=\left[\lambda\left(\frac{\alpha}{3}+\frac{\beta}{2}\right)+1\right]x+\lambda\left(\frac{\alpha}{4}+\frac{\beta}{3}\right),$$

$$\begin{cases}\left(1-\dfrac{\lambda}{3}\right)\alpha-\dfrac{\lambda}{2}\beta=1,\\[2mm] -\dfrac{\lambda}{4}\alpha+\left(1-\dfrac{\lambda}{3}\right)\beta=0;\end{cases}$$

$$\begin{cases}\alpha=\dfrac{24\lambda-72}{\lambda^2+48\lambda-72},\\[2mm] \beta=\dfrac{-18\lambda}{\lambda^2+48\lambda-72};\end{cases}$$

$$\varphi(x)=\frac{(24\lambda-72)x-18\lambda}{\lambda^2+48\lambda-72}.$$

这与前面的结果一致.

7.3 求下列积分方程 $\varphi(x)=x^2+\lambda\int_0^1(xy+y^2)\varphi(y)\mathrm{d}y$ 的解.

解 (a) 由于 $f(x)=x^2$ 不能用齐次积分方程的本征函数作线性展开,因此不能直接采用 Hilbert-Schmidt 方法求解. 为此,用 $\varphi(x)=x^2+\psi(x)$ 作变换,得到关于 $\psi(x)$ 的下列积分方程

$$\psi(x)=\lambda\left(\frac{1}{4}x+\frac{1}{5}\right)+\lambda\int_0^1(xy+y^2)\psi(y)\mathrm{d}y.$$

这里 $f(x)=\lambda\left(\frac{1}{4}x+\frac{1}{5}\right)=\lambda\int_0^1(xy+y^2)y^2\mathrm{d}y$,可用其齐次积分方程的本征函数(形式同前题)表达为

$$f(x)=\lambda\left(\frac{1}{4}x+\frac{1}{5}\right)=\lambda\left[\left(\frac{\sqrt{2}}{16}+\frac{1}{10}\right)\varphi^{(1)}(x)+\left(\frac{\sqrt{2}}{16}-\frac{1}{10}\right)\varphi^{(2)}(x)\right].$$

于是,容易按 Hilbert-Schmidt 方法求得

$$\varphi(x)=x^2-\frac{12\lambda(\lambda+15)x-3\lambda(\lambda-48)}{10(\lambda^2+48\lambda-72)}.$$

(b) 试验法. 取 $\varphi(x)=x^2+\alpha x+\beta$,则有

$$\begin{cases}\left(1-\dfrac{\lambda}{3}\right)\alpha-\dfrac{\lambda}{2}\beta=\dfrac{\lambda}{4},\\[2mm]-\dfrac{\lambda}{4}\alpha+\left(1-\dfrac{\lambda}{3}\right)\beta=\dfrac{\lambda}{5};\end{cases}$$

同样很容易求得上述结果.

(c) Fredholm 解法. 由 7.2.4 节例子已经求出 Fredholm 行列式和初余子式分别为

$$D(\lambda)=1-\frac{2}{3}\lambda-\frac{1}{72}\lambda^2,$$

$$D(x,y\mid\lambda)=\lambda(xy+y^2)+\lambda^2\left(\frac{1}{2}xy^2-\frac{1}{3}xy-\frac{1}{3}y^2+\frac{1}{4}y\right);$$

容易求得

$$\begin{aligned}\varphi(x)&=x^2+\frac{1}{D(\lambda)}\int_0^1 D(x,y\mid\lambda)y^2\mathrm{d}y\\&=x^2-\frac{12\lambda(\lambda+15)x-3\lambda(\lambda-48)}{10(\lambda^2+48\lambda-72)}.\end{aligned}$$

7.4 试证明:若 λ_0 为对称核 $K(x,y)$ 的一个本征值,则积分方程

$$\varphi(x)=f(x)+\lambda_0\int_a^b K(x,y)\varphi(y)\mathrm{d}y$$

无解,除非 $f(x)$ 与对应于 λ_0 的所有本征函数均正交.

证 首先,因为 $D(\lambda_0)=0$,故积分方程一般无解. 但是,若 $f(x)$ 与对应于 λ_0

的本征函数 $\varphi^{(0)}(x)$ 正交，则可证明其肯定有解.（为简单起见，设 λ_0 为单根，对重根情况可类似地证明.）

注意到 $K(x,y)$ 可展开为

$$K(x,y)=\bar{K}(x,y)+\frac{1}{\lambda_0}\varphi^{(0)}(x)\varphi^{(0)}(y),$$

$$\bar{K}(x,y)=\sum{}'\frac{1}{\lambda_n}\varphi^{(n)}(x)\varphi^{(n)}(y),$$

其中 $\sum'$ 表示对与 $\varphi^{(0)}$ 正交的所有 $\varphi^{(n)}$ 之项求和. 现在再令

$$\bar{\varphi}(x)=\varphi(x)-z_0\varphi^{(0)}(x),\quad z_0\equiv\int_a^b\varphi(x)\varphi^{(0)}(x)\mathrm{d}x;$$

于是有

$$\begin{aligned}\lambda_0\int_a^b K(x,y)\varphi(y)\mathrm{d}y&=\lambda_0\int_a^b\bar{K}(x,y)\varphi(y)\mathrm{d}y+z_0\varphi^{(0)}(x)\\&=\lambda_0\int_a^b\bar{K}(x,y)\bar{\varphi}(y)\mathrm{d}y+z_0\varphi^{(0)}(x).\end{aligned}$$

因而原积分方程变为

$$\bar{\varphi}(x)=f(x)+\lambda_0\int_a^b\bar{K}(x,y)\bar{\varphi}(y)\mathrm{d}y.$$

现在 λ_0 并非 $\bar{K}(x,y)$ 的本征值，即 $\bar{D}(\lambda_0)\neq0$，故一般有解. 若用 $\int_a^b\cdots\varphi^{(0)}(x)\mathrm{d}x$ 作用于上述方程两边，得到

$$\int_a^b f(x)\varphi^{(0)}(x)\mathrm{d}x=0.$$

即要求 $f(x)$ 与 $\varphi^{(0)}(x)$ 正交. 也就是说，$f(x)$ 处于 $\varphi(x)$ 的子空间，即与 $\bar{\varphi}(x)$ 处于同一函数空间.

7.5　试证明具有对称核的第一类积分方程

$$f(x)=\int_a^b K(x,y)\varphi(y)\mathrm{d}y$$

的一个解是

$$\varphi(x)=\sum_n a_n\lambda_n\varphi^{(n)}(x),$$

条件是该级数一致收敛；这里 $\lambda_n,\varphi^{(n)}(x)$ 是 $K(x,y)$ 的本征值和本征函数，而 a_n 是

$$f(x)=\sum_n a_n\varphi^{(n)}(x)$$

这一展开式中的系数.

证　若 $\varphi(x)=\sum_n a_n\lambda_n\varphi^{(n)}(x)$ 一致收敛，可将其代入积分方程右边进行逐项积分得到

$$\begin{aligned}\int_a^b K(x,y)\varphi(y)\mathrm{d}y &= \sum_n a_n\lambda_n\int_a^b K(x,y)\varphi^{(n)}(y)\mathrm{d}y \\ &= \sum_n a_n\lambda_n \cdot \frac{1}{\lambda_n}\varphi^{(n)}(x) \\ &= \sum_n a_n\varphi^{(n)}(x) = f(x),\end{aligned}$$

得证.

第 8 章　函 数 空 间

8.1 引　　言

8.1.1 基本概念

以函数作为点所构成的无穷维线性空间，称为**函数空间**.

考虑定义于区间(a,b)上的任意二函数，$f(x)$和$g(x)$，可仿照矢量空间中矢量标积的概念，引进

$$(f,g)=\int_a^b f(x)g(x)\mathrm{d}x$$

作为函数空间中**二函数的标积或内积**. 同时可定义**函数 $f(x)$ 的长度** $|f|$ 为

$$|f|=\sqrt{\int_a^b [f(x)]^2\mathrm{d}x}=\sqrt{(f,f)},$$

二函数 $f(x)$ 和 $g(x)$ 之间的距离 $|f-g|$ 为

$$|f-g|=\sqrt{\int_a^b [f(x)-g(x)]^2\mathrm{d}x}=\sqrt{(f-g,f-g)},$$

以及**二函数之间夹角**的余弦为

$$\cos\theta=\frac{(f,g)}{\sqrt{(f,f)}\sqrt{(g,g)}}.$$

8.1.2 Schwarz 不等式

因为三角形两边之和大于第三边，

$$|\boldsymbol{f}|+|\boldsymbol{g}|\geqslant|\boldsymbol{f}+\boldsymbol{g}|,$$

或

$$\sqrt{(\boldsymbol{f},\boldsymbol{f})}+\sqrt{(\boldsymbol{g},\boldsymbol{g})}\geqslant\sqrt{(\boldsymbol{f}+\boldsymbol{g},\boldsymbol{f}+\boldsymbol{g})},$$

两边平方后得

$$\begin{aligned}&(\boldsymbol{f},\boldsymbol{f})+(\boldsymbol{g},\boldsymbol{g})+2\sqrt{(\boldsymbol{f},\boldsymbol{f})(\boldsymbol{g},\boldsymbol{g})}\\&\quad\geqslant(\boldsymbol{f}+\boldsymbol{g},\boldsymbol{f}+\boldsymbol{g})=(\boldsymbol{f},\boldsymbol{f})+(\boldsymbol{g},\boldsymbol{g})+2(\boldsymbol{f},\boldsymbol{g});\end{aligned}$$

于是，消去相同项后两边再平方得

$$(\boldsymbol{f},\boldsymbol{f})(\boldsymbol{g},\boldsymbol{g})\geqslant\{(\boldsymbol{f},\boldsymbol{g})\}^2.$$

推广至现在的函数空间，容易得到

$$\int_a^b f(x)f(x)\mathrm{d}x\int_a^b g(x)g(x)\mathrm{d}x \geqslant \left\{\int_a^b f(x)g(x)\mathrm{d}x\right\}^2.$$

或者，将下列显然结果

$$\int_a^b\left\{\int_a^b [f(x)g(y)-f(y)g(x)]^2\mathrm{d}y\right\}\mathrm{d}x \geqslant 0$$

展开后，容易得到

$$\begin{aligned}
0 &\leqslant \int_a^b\left\{\int_a^b [f^2(x)g^2(y)+f^2(y)g^2(x)-2f(x)g(x)f(y)g(y)]\mathrm{d}y\right\}\mathrm{d}x \\
&= \int_a^b f^2(x)\mathrm{d}x\int_a^b g^2(y)\mathrm{d}y + \int_a^b g^2(x)\mathrm{d}x\int_a^b f^2(y)\mathrm{d}y \\
&\quad - 2\int_a^b f(x)g(x)\mathrm{d}x\int_a^b f(y)g(y)\mathrm{d}y \\
&= 2\left\{\int_a^b f(x)f(x)\mathrm{d}x\int_a^b g(x)g(x)\mathrm{d}x - \left[\int_a^b f(x)g(x)\mathrm{d}x\right]^2\right\},
\end{aligned}$$

即

$$\int_a^b f(x)f(x)\mathrm{d}x\int_a^b g(x)g(x)\mathrm{d}x \geqslant \left|\int_a^b f(x)g(x)\mathrm{d}x\right|^2.$$

以上结果称为 **Schwarz 不等式**.

由于此不等式，表明前面所定义函数空间中两函数夹角余弦的正确性.

8.1.3 备注

应注意：(i) 函数的定义区间(a,b)也可以为无限，如 $a=-\infty$, $b=\infty$；后面将不另加说明. (ii) 对于复值函数，(f,g)的积分中之 f 应取其复共轭 f^*. (iii) 推广至多元函数的情况是很容易的.

另一方面，矢量空间中常取正交归一矢量集合为其基；与此类似，函数空间中则常取正交归一函数系为其基. 在 7.4.4 节曾讨论过正交归一本征函数系，下面将首先一般地论述正交归一函数系及其性质.

8.2 正交归一函数系

8.2.1 定义

一组连续实函数 $\varphi^{(1)}(x)$, $\varphi^{(2)}(x)$, $\varphi^{(3)}(x)$, …，若它们在区间(a,b)上满足条件：

$$\int_a^b \varphi^{(n)}(x)\varphi^{(m)}(x)\mathrm{d}x = 0 \quad (n\neq m),$$

则称之为互相**正交**的($\varphi^{(n)} \perp \varphi^{(m)}$);另外,若满足条件

$$\int_a^b \varphi^{(n)}(x)\varphi^{(n)}(x)\mathrm{d}x = 1,$$

则称之为**归一化**的($\varphi^{(n)}$的"长"$=1$);或合记成

$$\int_a^b \varphi^{(n)}(x)\varphi^{(m)}(x)\mathrm{d}x = \delta_{nm},$$

并称这样一组函数为**正交归一函数系**.

正交归一函数系构成函数空间的基[函数系],它与矢量空间的基[矢系]类似,例如,上述条件相当于

$$\boldsymbol{\varphi}^{(n)} \cdot \boldsymbol{\varphi}^{(m)} = \sum_i \varphi_i^{(n)} \varphi_i^{(m)} = \delta_{nm}.$$

它还具有与之相应的一些性质,如线性无关性,完备性等. 下面分别予以讨论.

8.2.2 线性无关性

正交归一函数系是**线性无关**的;即,如果

$$\sum c^{(n)} \varphi^{(n)}(x) = c^{(1)}\varphi^{(1)}(x) + \cdots + c^{(n)}\varphi^{(n)}(x) + \cdots = 0,$$

则

$$c^{(1)} = c^{(2)} = \cdots = c^{(n)} = \cdots = 0.$$

证 用 $\int_a^b \cdots \varphi^{(n)}(x)\mathrm{d}x$ 作用于前式,得到

$$0 = \int_a^b (c^{(1)}\varphi^{(1)}(x) + \cdots + c^{(n)}\varphi^{(n)}(x) + \cdots)\varphi^{(n)}(x)\mathrm{d}x = c^{(n)},$$

这里应用了 $\varphi^{(m)}(x)$的正交归一性质,因而得到上式,即 $\varphi^{(1)}(x),\cdots,\varphi^{(n)}(x),\cdots$是线性无关的.

8.2.3 完备性

函数空间中定义于区间(a,b)上的任意函数 $f(x)$均可用该空间的基函数 $\varphi^{(n)}(x)$的线性组合来表示,即

$$f(x) = \sum c^{(n)} \varphi^{(n)}(x),$$

这里 $c^{(n)}$是 $f(x)$在 $\varphi^{(n)}(x)$方向的投影,即 $f(x)$与 $\varphi^{(n)}(x)$的标积,

$$c^{(n)} = \int_a^b \varphi^{(n)}(x) f(x)\mathrm{d}x.$$

此性质称为正交归一函数系的**完备性**.

8.3 Fourier 级数

Fourier级数理论是以三角函数系为基的函数空间的典型例子.

8.3.1 三角函数系

下列函数系：

$$\varphi^{(0)}(x)=1/\sqrt{2\pi},$$

$$\varphi^{(1)}(x)=\frac{1}{\sqrt{\pi}}\sin x,\quad \varphi^{(2)}(x)=\frac{1}{\sqrt{\pi}}\cos x,$$

$$\varphi^{(3)}(x)=\frac{1}{\sqrt{\pi}}\sin 2x,\quad \varphi^{(4)}(x)=\frac{1}{\sqrt{\pi}}\cos 2x,$$

$$\vdots \qquad\qquad \vdots$$

$$\varphi^{(2r-1)}(x)=\frac{1}{\sqrt{\pi}}\sin rx,\quad \varphi^{(2r)}(x)=\frac{1}{\sqrt{\pi}}\cos rx,$$

等等；是在区间$[-\pi,\pi]$上的正交归一函数系，称为**三角函数系**. 它们是归一化的，因为

$$\int_{-\pi}^{\pi}\varphi^{(0)}(x)\varphi^{(0)}(x)\mathrm{d}x=\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{d}x=1,$$

$$\int_{-\pi}^{\pi}\varphi^{(2r-1)}(x)\varphi^{(2r-1)}(x)\mathrm{d}x=\frac{1}{\pi}\int_{-\pi}^{\pi}\sin^2 rx\,\mathrm{d}x=\frac{1}{2\pi}\int_{-\pi}^{\pi}(1-\cos 2rx)\mathrm{d}x=1,$$

$$\int_{-\pi}^{\pi}\varphi^{(2r)}(x)\varphi^{(2r)}(x)\mathrm{d}x=\frac{1}{\pi}\int_{-\pi}^{\pi}\cos^2 rx\,\mathrm{d}x=\frac{1}{2\pi}\int_{-\pi}^{\pi}(1+\cos 2rx)\mathrm{d}x=1;$$

它们是相互正交的，因为

$$\int_{-\pi}^{\pi}\varphi^{(0)}(x)\varphi^{(2r-1)}(x)\mathrm{d}x=\frac{1}{\sqrt{2}\pi}\int_{-\pi}^{\pi}\sin rx\,\mathrm{d}x=0\quad(r\geqslant 1),$$

$$\int_{-\pi}^{\pi}\varphi^{(0)}(x)\varphi^{(2r)}(x)\mathrm{d}x=\frac{1}{\sqrt{2}\pi}\int_{-\pi}^{\pi}\cos rx\,\mathrm{d}x=0\quad(r\geqslant 1),$$

$$\frac{1}{\pi}\int_{-\pi}^{\pi}\sin rx\ \sin sx\,\mathrm{d}x=\frac{1}{2\pi}\int_{-\pi}^{\pi}[\cos(r-s)x-\cos(r+s)x]\mathrm{d}x=0\quad(r\neq s),$$

$$\frac{1}{\pi}\int_{-\pi}^{\pi}\cos rx\ \cos sx\,\mathrm{d}x=\frac{1}{2\pi}\int_{-\pi}^{\pi}[\cos(r-s)x+\cos(r+s)x]\mathrm{d}x=0\quad(r\neq s),$$

$$\frac{1}{\pi}\int_{-\pi}^{\pi}\sin rx\ \cos sx\,\mathrm{d}x=\frac{1}{2\pi}\int_{-\pi}^{\pi}[\sin(r+s)x+\sin(r-s)x]\mathrm{d}x=0.$$

8.3.2 函数的 Fourier 级数展开

定义于$(-\pi,\pi)$上的函数$f(x)$可用三角函数系表示为

$$\begin{aligned}f(x)&=\frac{1}{2}a_0+(a_1\cos x+b_1\sin x)+(a_2\cos 2x+b_2\sin 2x)+\cdots\\&=\frac{1}{2}a_0+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx),\end{aligned}$$

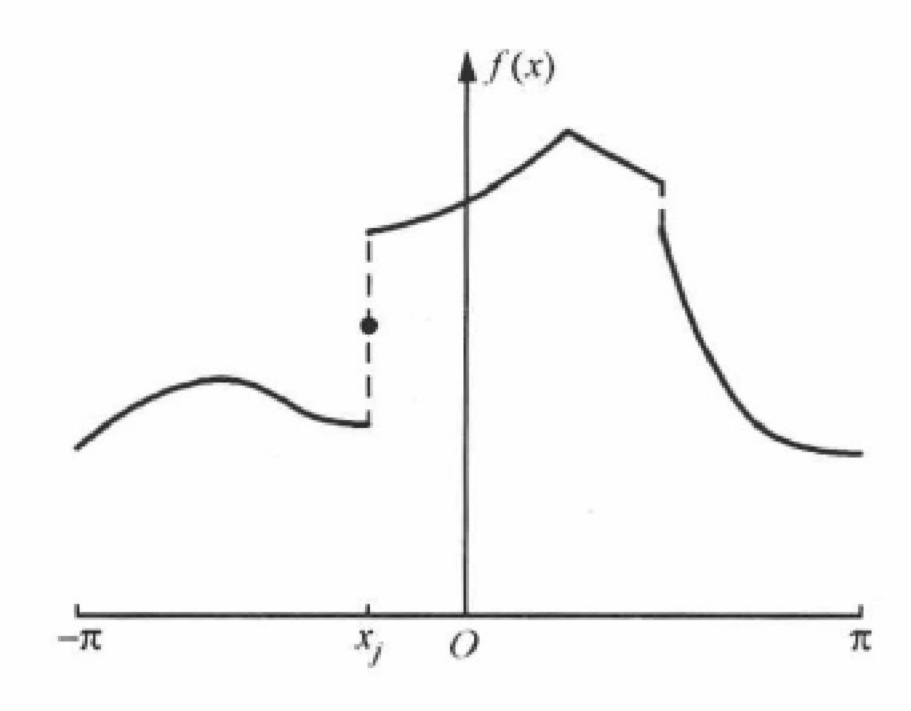
其中

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,\mathrm{d}x, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,\mathrm{d}x \quad (n = 0,1,2,\cdots);$$

这样的展开称为 $f(x)$的 **Fourier 级数**，而 a_n 和 b_n 则称为 $f(x)$的 **Fourier 系数**.

应该指出的是：这里的"="须证明；大致说[①]，若 $f(x)$是具有限跃变的分段连续函数(如下图)，则其 Fourier 级数在连续点 x 收敛于 $f(x)$，而在跃变点 x_j 则收敛于

$$\frac{1}{2}[f(x_j+) + f(x_j-)].$$



其次应指出，若函数 $f(x)$定义于任意区间(a,b)，可通过引进新变量

$$y = \frac{2\pi}{b-a}\left[x - \frac{1}{2}(b+a)\right],$$

并定义新函数 $h(y)=f(x)$，则 $h(y)$ 定义于区间$(-\pi,\pi)$上；对此可按前述方式进行 Fourier 级数展开. 最后再变换回原变量 x 和原函数 $f(x)$，则得任意区间上的 Fourier 级数.

8.3.3 Parseval 公式

定义于$(-\pi,\pi)$上的函数 $f(x)$用三角函数系展开为 Fourier 级数时，有下列 **Parseval 公式**：

$$\int_{-\pi}^{\pi} f(x)f(x)\mathrm{d}x = \pi\left\{\frac{1}{2}a_0^2 + \sum_{n=1}^{\infty}(a_n^2 + b_n^2)\right\}.$$

定义于(a,b)上的函数 $f(x)$，$g(x)$可用正交归一函数系 $\varphi_0(x)$，$\varphi_1(x)$，$\varphi_2(x)$，…，$\varphi_n(x)$，…展开为广义 Fourier 级数

$$f(x) = \sum_{n=0}^{\infty} c^{(n)}\varphi^{(n)}(x), \quad c^{(n)} = \int_a^b \varphi^{(n)}(x)f(x)\mathrm{d}x,$$

① E. T. Whittaker, G. N. Watson, *Modern Analysis*, §9.42.

$$g(x)=\sum_{n=0}^{\infty}d^{(n)}\varphi^{(n)}(x),\quad d^{(n)}=\int_a^b\varphi^{(n)}(x)g(x)\mathrm{d}x;$$

同样有 **Parseval** 公式：

$$\int_a^b f(x)f(x)\mathrm{d}x=\sum_{n=0}^{\infty}c^{(n)2},$$

以及

$$\int_a^b f(x)g(x)\mathrm{d}x=\sum_{n=0}^{\infty}c^{(n)}d^{(n)};$$

它们类似于矢量空间中用正交基 $\boldsymbol{\xi}^{(1)},\cdots,\boldsymbol{\xi}^{(n)}$ 展开时的标积，即，若

$$\boldsymbol{f}=f_1\boldsymbol{\xi}^{(1)}+\cdots+f_n\boldsymbol{\xi}^{n},\quad \boldsymbol{g}=g_1\boldsymbol{\xi}^{(1)}+\cdots+g_n\boldsymbol{\xi}^{(n)},$$

则有

$$(\boldsymbol{f},\boldsymbol{f})=f_1^2+\cdots+f_n^2,\quad (\boldsymbol{f},\boldsymbol{g})=f_1g_1+\cdots+f_ng_n.$$

8.3.4　应用例子

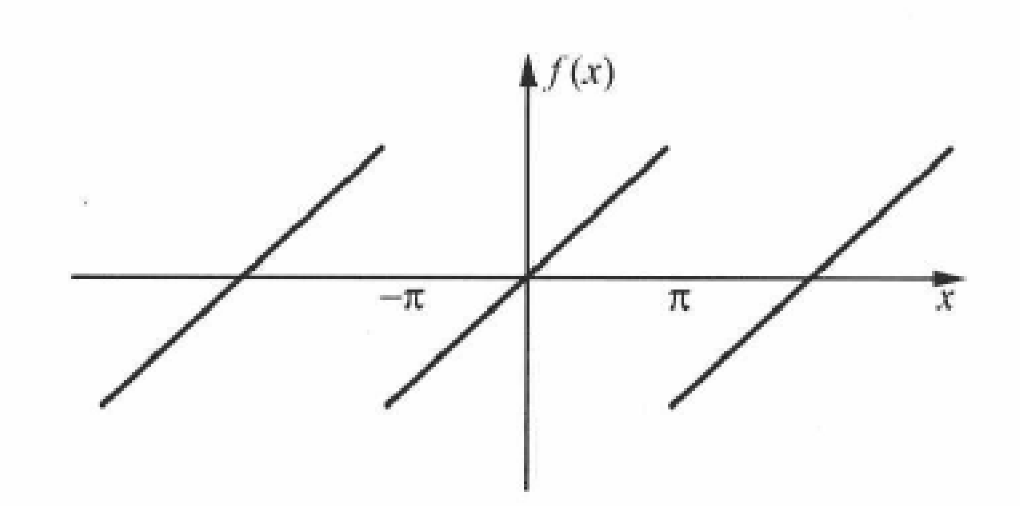
函数 $f(x)$ 的 Fourier 级数是

$$\bar{f}(x)=\frac{1}{2}a_0+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx),$$

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\,\mathrm{d}x,\quad b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nx\,\mathrm{d}x\quad(n=0,1,2,\cdots).$$

应注意，级数展开式给出的 $\bar{f}(x)$ 是周期为 2π 的周期函数，而确定 Fourier 系数 a_n,b_n 中的原函数 $f(x)$ 只需在区间 $(-\pi,\pi)$ 予以定义即可.

例 1　如下图



$$f(x)=x\quad(-\pi<x<\pi).$$

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}x\cos nx\,\mathrm{d}x=0,$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}x\sin nx\,\mathrm{d}x=-\frac{\mathrm{d}}{\mathrm{d}n}\left\{\frac{1}{\pi}\int_{-\pi}^{\pi}\cos nx\,\mathrm{d}x\right\}$$

$$=\frac{\mathrm{d}}{\mathrm{d}n}\left\{-\frac{\sin nx}{n\pi}\right\}\Bigg|_{-\pi}^{\pi}=-\frac{x\cos nx}{n\pi}\Bigg|_{-\pi}^{\pi}$$

$$=-\frac{2\pi\cos n\pi}{n\pi}=-\frac{2}{n}(-1)^n,$$

$$\text{或 } b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}x\sin nx\,\mathrm{d}x=-\frac{1}{n\pi}\int_{-\pi}^{\pi}x\mathrm{d}\cos nx$$

$$=-\frac{1}{n\pi}\left\{x\cos nx\Big|_{-\pi}^{\pi}-\int_{-\pi}^{\pi}\cos nx\,\mathrm{d}x\right\}$$

$$=-\frac{1}{n\pi}2\pi\cos n\pi=-\frac{2}{n}(-1)^n.$$

$$\therefore\quad a_n=0,\quad b_n=(-1)^{n+1}\frac{2}{n};$$

于是，$f(x)$的 Fourier 级数 $\bar{f}(x)$是

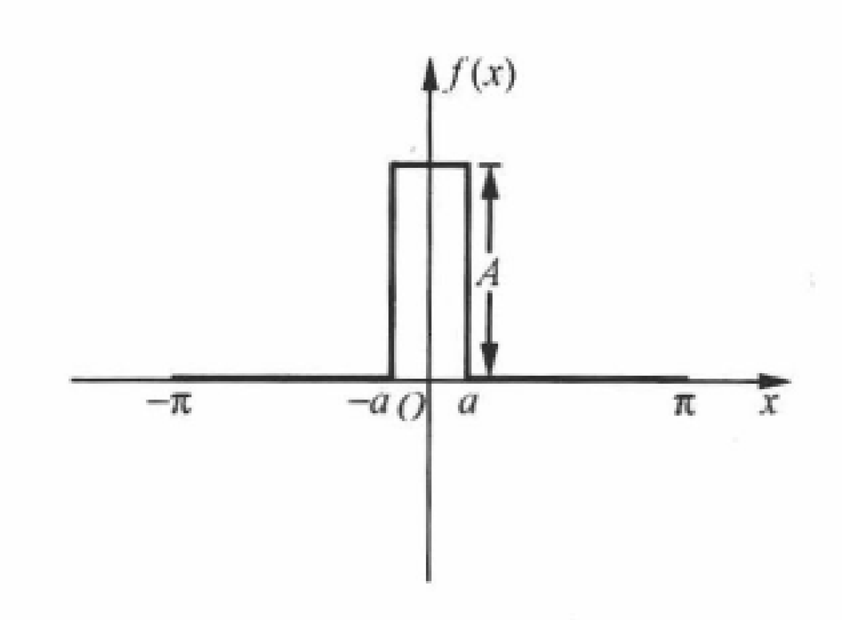
$$\bar{f}(x)=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{2}{n}\sin(nx);$$

这样的级数显然是收敛的.

另外注意到，在跃变点，如 $x_j=\pm\pi$，级数给出 $\bar{f}(x_j)=0$. 也就是说，它的确收敛于跃变点的平均值，即，

$$\bar{f}(\pm\pi)=\frac{1}{2}[f(\pm\pi+0)+f(\pm\pi-0)]=\frac{1}{2}[-\pi+\pi]=0.$$

例 2 如下图



$$f(x)=\begin{cases}0, & -\pi<x<-a,\\ A=\sqrt{\dfrac{\pi}{2a}}, & -a<x<a,\\ 0, & a<x<\pi,\end{cases}$$

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\,\mathrm{d}x=\frac{A}{\pi}\int_{-a}^{a}\cos nx\,\mathrm{d}x$$

$$=\frac{A}{\pi}\frac{\sin nx}{n}\Big|_{-a}^{a}=\frac{2A}{n\pi}\sin na,$$

$$a_0=\frac{2Aa}{\pi},$$

$$b_n = 0.$$

$$\therefore \quad f(x) = \frac{Aa}{\pi} + \sum_{n=1}^{\infty} \frac{2A}{n\pi} \sin na \cos nx.$$

因为

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2 \mathrm{d}x = \frac{1}{\pi}\int_{-a}^{a} A^2 \mathrm{d}x = \frac{2a}{\pi} A^2,$$

若选 $A=\sqrt{\frac{\pi}{2a}}$,则上式结果为 1. 于是,由 Parseral 公式

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2 \mathrm{d}x = \frac{1}{2}a_0^2 + \sum_{n=1}^{\infty}(a_n^2 + b_n^2),$$

和上面求得的 Fourier 系数,得到

$$\frac{2aA^2}{\pi} = \frac{2a^2A^2}{\pi^2} + \sum_{n=1}^{\infty} \frac{4A^2}{n^2\pi^2} \sin^2 na.$$

消去 A^2 并注意到当 $a \neq 0$ 时,则有

$$1 = \frac{a}{\pi} + \sum_{n=1}^{\infty} \frac{2}{n^2 \pi a} \sin^2 na,$$

$$\frac{\pi}{a} = 1 + \sum_{n=1}^{\infty} \frac{2}{n^2 a^2} \sin^2 na = \sum_{n=-\infty}^{\infty} \frac{\sin^2 na}{n^2 a^2},$$

或

$$\sum_{n=-\infty}^{\infty} \frac{\sin^2 na}{n^2 a} = \pi.$$

现在令 $x = na$, $\Delta x = a\Delta n(\Delta n = 1, \Delta x = a)$,当 $a \to 0$ 时,得到下列结果:

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} \mathrm{d}x = \pi.$$

8.3.5 Fourier 级数的复数形式

因为 $\mathrm{e}^{\pm inx} = \cos nx \pm \mathrm{i} \sin nx$,故有

$$\cos nx = \frac{1}{2}(\mathrm{e}^{inx} + \mathrm{e}^{-inx}),$$

$$\sin nx = \frac{1}{2\mathrm{i}}(\mathrm{e}^{inx} - \mathrm{e}^{-inx}).$$

若令

$$\frac{1}{2}(a_n - \mathrm{i}b_n) = c_n, \quad \frac{1}{2}(a_n + \mathrm{i}b_n) = c_{-n},$$

则 $f(x)$ 的 Fourier 级数

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$$

可以写成下列复数形式

$$f(x)=\sum_{n=-\infty}^{\infty}c_n\mathrm{e}^{\mathrm{i}nx};$$

同时,Fourier 系数 c_n 可求出为

$$c_n=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x.$$

8.3.6 二维和三维空间的情形

由上述 Fourier 级数的复数形式,不难推广到二维和三维空间的情形.

二维空间的 Fourier 级数为

$$f(x,y)=\sum_{l,m=-\infty}^{\infty}c_{lm}\mathrm{e}^{\mathrm{i}lx+\mathrm{i}my},$$

$$c_{lm}=\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f(x,y)\mathrm{e}^{-\mathrm{i}lx-\mathrm{i}my}\,\mathrm{d}x\mathrm{d}y.$$

三维空间的 Fourier 级数为

$$f(x,y,z)=\sum_{l,m,n=-\infty}^{\infty}c_{lmn}\mathrm{e}^{\mathrm{i}lx+\mathrm{i}my+\mathrm{i}nz},$$

$$c_{lmn}=\frac{1}{(2\pi)^3}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f(x,y,z)\mathrm{e}^{-\mathrm{i}lx-\mathrm{i}my-\mathrm{i}nz}\,\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

8.4 Fourier 变换

8.4.1 Fourier 积分

$f(x)$的 Fourier 级数对 x 而言周期是 $2\pi(-\pi\to\pi)$,若令 $x\to\frac{x}{N}$,则对后来的 x 而言周期是 $2N\pi(-N\pi\to N\pi)$;然后再令 $N\to\infty$,将使 Fourier 级数推广为 Fourier 积分,其周期从有限变为无限. 现推导如下:

当 $f(x)$的周期是 $2\lambda\pi$ 时的 Fourier 级数为

$$f(x)=\frac{1}{2}a_0+\sum_{n=1}^{\infty}\left(\lambda a_n\cos\frac{nx}{\lambda}+\lambda b_n\sin\frac{nx}{\lambda}\right)\frac{\Delta n}{\lambda},$$

$$\lambda a_m=\frac{1}{\pi}\int_{-\pi\lambda}^{\pi\lambda}f(x)\cos\frac{mx}{\lambda}\mathrm{d}x,$$

$$\lambda b_m=\frac{1}{\pi}\int_{-\pi\lambda}^{\pi\lambda}f(x)\sin\frac{mx}{\lambda}\mathrm{d}x,$$

这里 $\Delta n=n-(n-1)=1$. 当 $\lambda\to\infty$时,令$\frac{n}{\lambda}\to u$, $\frac{\Delta n}{\lambda}\to\mathrm{d}u$, $\lambda a_n\to a(u)$, $\lambda b_n\to b(u)$,

而 $\sum_{n=0}^{\infty}(\quad)\frac{\Delta n}{\lambda}\to\int_0^{\infty}(\quad)\mathrm{d}u$，于是得到

$$f(x)=\int_0^{\infty}(a(u)\cos ux+b(u)\sin ux)\mathrm{d}u,$$

$$a(u)=\frac{1}{\pi}\int_{-\infty}^{\infty}f(x)\cos ux\,\mathrm{d}x,$$

$$b(u)=\frac{1}{\pi}\int_{-\infty}^{\infty}f(x)\sin ux\,\mathrm{d}x,$$

这个结果称为 $f(x)$ 的 **Fourier 积分.**

当 $f(x)$ 是偶或奇函数时，Fourier 积分也得到简化. 具体地说，对于偶函数 $f(x)=f(-x)$，有 $b(u)=0$；而对于奇函数 $f(x)=-f(-x)$，则 $a(u)=0$. 这与 Fourier 级数的情况一致.

还可将 $a(u)$，$b(u)$ 代入 $f(x)$ 的表达式中，容易得到

$$f(x)=\frac{1}{\pi}\int_0^{\infty}\mathrm{d}u\int_{-\infty}^{\infty}f(y)\cos u(x-y)\mathrm{d}y,$$

有时将上述结果称为 **Fourier 积分定理.**

8.4.2　Fourier 积分的复数形式

因为 $\mathrm{e}^{\pm i\theta}=\cos\theta\pm \mathrm{i}\sin\theta$，故有

$$\cos\theta=\frac{1}{2}(\mathrm{e}^{i\theta}+\mathrm{e}^{-i\theta}),$$

$$\sin\theta=\frac{1}{2\mathrm{i}}(\mathrm{e}^{i\theta}-\mathrm{e}^{-i\theta}).$$

于是 $f(x)$ 的 Fourier 积分可化为

$$\begin{aligned}f(x)&=\int_0^{\infty}\{a(u)\cos ux+b(u)\sin ux\}\mathrm{d}u\\&=\frac{1}{2}\int_0^{\infty}\{[a(u)-\mathrm{i}b(u)]\mathrm{e}^{\mathrm{i}ux}+[a(u)+\mathrm{i}b(u)]\mathrm{e}^{-\mathrm{i}ux}\}\mathrm{d}u\\&=\frac{1}{2}\int_{-\infty}^{0}[a(-u)-\mathrm{i}b(-u)]\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}u+\frac{1}{2}\int_0^{\infty}[a(u)+\mathrm{i}b(u)]\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}u.\end{aligned}$$

由 $a(u)$ 和 $b(u)$ 的表达式容易求得

$$a(u)+\mathrm{i}b(u)=a(-u)-\mathrm{i}b(-u)=\frac{1}{\pi}\int_{-\infty}^{\infty}f(x)\mathrm{e}^{\mathrm{i}ux}\mathrm{d}x\equiv\frac{1}{\pi}F(u).$$

最后将它代入上式，得到 Fourier 积分的复数形式为

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(u)\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}u,$$

$$F(u)=\int_{-\infty}^{\infty}f(x)\mathrm{e}^{\mathrm{i}ux}\mathrm{d}x.$$

还可将后式代入前式，得到

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}ux}\,\mathrm{d}u\int_{-\infty}^{\infty}f(y)\mathrm{e}^{\mathrm{i}uy}\,\mathrm{d}y$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}u\int_{-\infty}^{\infty}f(y)\mathrm{e}^{-\mathrm{i}u(x-y)}\,\mathrm{d}y,$$

有时将上述结果称为 **Fourier 积分定理.**

由于上一小节的最后一个积分中 $\int_{-\infty}^{\infty}f(y)\cos u(x-y)\mathrm{d}y$ 是 u 的偶函数，故该式对 u 的积分可从 $\int_{0}^{\infty}(\quad)\mathrm{d}u$ 变换为 $\frac{1}{2}\int_{-\infty}^{\infty}(\quad)\mathrm{d}u$；另外，由于 $\int_{-\infty}^{\infty}f(y)\sin u(x-y)\mathrm{d}y$ 是 u 的奇函数，后者对 $\int_{-\infty}^{\infty}(\quad)\mathrm{d}u$ 的积分为零. 于是，利用 $\mathrm{e}^{\mathrm{i}\theta}=\cos\theta+\mathrm{i}\sin\theta$，同样可将该 Fourier 积分改写为上列复数形式.

8.4.3 Fourier 变换与 Fourier 逆变换

上面得到的典型形式

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(u)\mathrm{e}^{-\mathrm{i}ux}\,\mathrm{d}u,$$

$$F(u)=\int_{-\infty}^{\infty}f(x)\mathrm{e}^{\mathrm{i}ux}\,\mathrm{d}x,$$

后者将函数 $f(x)$ 变换为 $F(u)$，可记做 $\mathscr{F}(f(x))=F(u)$，称为 $f(x)$ 的 **Fourier 变换**；而前者则将 $F(u)$ 变换回 $f(x)$，可记做 $\mathscr{F}^{-1}(F(u))=f(x)$，故称为 $F(u)$ 的 **Fourier 逆变换.**

应该指出，文献上对于 Fourier 变换的定义并不唯一，但可由下列形式：

$$f(x)=\frac{|\gamma|\,C}{2\pi}\int_{-\infty}^{\infty}F(u)\mathrm{e}^{-\mathrm{i}\gamma ux}\,\mathrm{d}u,$$

$$F(u)=\frac{1}{C}\int_{-\infty}^{\infty}f(x)\mathrm{e}^{\mathrm{i}\gamma ux}\,\mathrm{d}x$$

相联系，其中 C 和 γ 是常数. 这里取 $C=\gamma=1$ 是目前物理上常用的形式. 有时取 $C=\sqrt{2\pi}$，$\gamma=1$，则是为了使之具有较对称形式. 其他定义则取 $C=1$，$C=\sqrt{2\pi}$，$C=2\pi$ 与 $\gamma=\pm1$，$\gamma=\pm2\pi$ 等的各种组合.

8.4.4 Parseval 公式

设有函数 $f(x)$ 和 $g(x)$，其 Fourier 积分分别为

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(u)\mathrm{e}^{-\mathrm{i}ux}\,\mathrm{d}u,$$

$$g(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}G(u)\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}u;$$

以及对前者两边取其复共轭得

$$f^*(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}F^*(u)\mathrm{e}^{\mathrm{i}ux}\mathrm{d}u=\frac{1}{2\pi}\int_{-\infty}^{\infty}F^*(-u)\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}u.$$

应该注意，$\mathscr{F}(f^*(x))=F^*(-u)$而$F^*(u)=[\mathscr{F}(f(x))]^*$.

现在利用上面的变换来计算下列积分：

$$\begin{aligned}\int_{-\infty}^{\infty}f(x)G(x)\mathrm{d}x&=\int_{-\infty}^{\infty}G(x)\mathrm{d}x\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty}F(u)\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}u\\&=\int_{-\infty}^{\infty}F(u)\mathrm{d}u\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty}G(x)\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}x\\&=\int_{-\infty}^{\infty}F(u)g(u)\mathrm{d}u;\end{aligned}$$

$$\begin{aligned}\int_{-\infty}^{\infty}f^*(x)g(x)\mathrm{d}x&=\int_{-\infty}^{\infty}\mathrm{d}x\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty}F^*(u)\mathrm{e}^{\mathrm{i}ux}\mathrm{d}u\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty}G(v)\mathrm{e}^{-\mathrm{i}vx}\mathrm{d}v\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}F^*(u)\mathrm{d}u\int_{-\infty}^{\infty}G(v)\mathrm{d}v\cdot\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(u-v)x}\mathrm{d}x\right]\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}F^*(u)\mathrm{d}u\int_{-\infty}^{\infty}G(v)\delta(u-v)\mathrm{d}v\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}F^*(u)G(u)\mathrm{d}u;\end{aligned}$$

这里假定积分次序可颠倒，后一积分计算中还应用了 Dirac δ 函数的性质. 最后结果重写如下：

$$\int_{-\infty}^{\infty}f(x)G(x)\mathrm{d}x=\int_{-\infty}^{\infty}F(u)g(u)\mathrm{d}u,$$

$$\int_{-\infty}^{\infty}f^*(x)g(x)\mathrm{d}x=\frac{1}{2\pi}\int_{-\infty}^{\infty}F^*(u)G(u)\mathrm{d}u.$$

特别是，在后一结果中若令 $g(x)=f(x)$，则得①

$$\int_{-\infty}^{\infty}|f(x)|^2\mathrm{d}x=\frac{1}{2\pi}\int_{-\infty}^{\infty}|F(u)|^2\mathrm{d}u.$$

这些结果称为 **Fourier 积分的 Parseval 公式**，它们与 Fourier 级数的 Parseval 公式完全类似.

利用这些公式可将许多难积的积分计算出来. 比如，若某个积分不容易积出，

① 此结果通常称为 **Plancherel 定理**.

则可通过适当划分写成公式中左边的一种形式，然后从 Fourier 变换表[①]上找出其对应 Fourier 变换，再对右边形式的积分进行计算，往往可以容易地积出.

8.4.5 卷积定理

函数 f 和 g 的**卷积** $f * g$ 定义为

$$f * g(x) = \int_{-\infty}^{\infty} f(s) g(x-s) \mathrm{d}s,$$

它是 x 的函数.

现在来求卷积的 Fourier 变换：

$$\begin{aligned}\mathscr{F}(f * g(x)) &= \int_{-\infty}^{\infty} f * g(x) \mathrm{e}^{\mathrm{i}ux} \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(s) g(x-s) \mathrm{d}s\right] \mathrm{e}^{\mathrm{i}ux} \mathrm{d}x \\ &= \int_{-\infty}^{\infty} f(s) \mathrm{e}^{\mathrm{i}us} \mathrm{d}s \left[\int_{-\infty}^{\infty} g(x-s) \mathrm{e}^{\mathrm{i}u(x-s)} \mathrm{d}(x-s)\right] \\ &= F(u) \cdot G(u),\end{aligned}$$

这里假定积分次序可颠倒.

这个结果表明：两函数 f 和 g 之卷积 $f * g$ 的 Fourier 变换 $\mathscr{F}(f * g)$ 等于 f 和 g 的 Fourier 变换 $\mathscr{F}(f)$ 和 $\mathscr{F}(g)$ 之积；即

$$\mathscr{F}(f * g) = \mathscr{F}(f) \cdot \mathscr{F}(g) = F \cdot G;$$

此结果称为 Fourier 变换的**卷积定理**.

8.4.6 应用例子

再将 Fourier 变换与 Fourier 逆变换写出如下：

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(u) \mathrm{e}^{-\mathrm{i}ux} \mathrm{d}u,$$

$$F(u) = \int_{-\infty}^{\infty} f(x) \mathrm{e}^{\mathrm{i}ux} \mathrm{d}x.$$

例 1 给定 $f(x) = \begin{cases} 1, & -a < x < +a; \\ 0, & -\infty < x < -a, +a < x < +\infty; \end{cases}$ 求 $F(u)$.

$$F(u) = \int_{-\infty}^{\infty} f(x) \mathrm{e}^{\mathrm{i}ux} \mathrm{d}x = \int_{-a}^{+a} \mathrm{e}^{\mathrm{i}ux} \mathrm{d}x = \left.\frac{\mathrm{e}^{\mathrm{i}ux}}{\mathrm{i}u}\right|_{-a}^{a} = 2\,\frac{\sin au}{u}.$$

---

① 例如，A. Erdelyi, W. Magnus, F. Oberhettinger, F. G. Tricomi, *Tables of Integral Transformations*, vol. I, McGraw-Hill, New York, 1954; p. 117ff. 该书采用的 Fourier 变换之定义相当于取 $C=1, \gamma=-1$；即

$$F(u) = \int_{-\infty}^{\infty} f(x) \mathrm{e}^{-\mathrm{i}ux} \mathrm{d}x.$$

另外，若给定 $g(x)=F(x)=\dfrac{2\sin ax}{x}$，求 $G(u)$. 首先，

$$G(u)=\int_{-\infty}^{\infty}F(x)\mathrm{e}^{\mathrm{i}ux}\,\mathrm{d}x=\int_{-\infty}^{\infty}F(v)\mathrm{e}^{-\mathrm{i}v(-u)}\,\mathrm{d}v=2\pi f(-u),$$

也就是说，$\mathscr{F}(\mathscr{F}(f(x)))=\mathscr{F}(F(x))=2\pi f(-u)$. 因此，

$$\begin{array}{cc}
f(x) & F(x)\\
F(x) & 2\pi f(-x)\\
\begin{cases}1, & |x|<a,\\ 0, & |x|>a\end{cases} & \dfrac{2\sin ax}{x}\\
\dfrac{2\sin ax}{x} & \begin{cases}1, & |x|<a,\\ 0, & |x|>a\end{cases}
\end{array}$$

同时，Parseval 公式：

$$\int_{-\infty}^{\infty}f(x)G(x)\,\mathrm{d}x=\int_{-\infty}^{\infty}F(u)g(u)\,\mathrm{d}u,$$

还可写成

$$\int_{-\infty}^{\infty}F(x)G(x)\,\mathrm{d}x=2\pi\int_{-\infty}^{\infty}f(-u)g(u)\,\mathrm{d}u.$$

现在若取

$$F(x)=\frac{2\sin ax}{x},\quad f(x)=\begin{cases}1, & |x|<a;\\ 0, & |x|>a;\end{cases}$$

$$G(x)=\frac{2\sin bx}{x},\quad g(x)=\begin{cases}1, & |x|<b;\\ 0, & |x|>b;\end{cases}$$

而且 $a>0,b>0$；则由 Parseval 公式的后一形式得到

$$\int_{-\infty}^{\infty}4\,\frac{\sin ax}{x}\,\frac{\sin bx}{x}\,\mathrm{d}x=2\pi\int_{-\min(a,b)}^{+\min(a,b)}1\,\mathrm{d}x=4\pi\min(a,b),$$

即

$$\int_{-\infty}^{\infty}\frac{\sin ax\sin bx}{x^2}\,\mathrm{d}x=\pi\min(a,b).$$

若取 $a=b=1$，则得

$$\int_{-\infty}^{\infty}\frac{\sin^2 x}{x^2}\,\mathrm{d}x=\pi,$$

这是 Fourier 级数的应用例子[8.3.4 节之例 2]中已经求得的结果.

另一方面，若取 $b<a$，令 a 固定而 b 变动，于是

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin ax\,\sin bx}{x^2}\,\mathrm{d}x=b,$$

$$\frac{\partial}{\partial b}\,\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin ax\,\sin bx}{x^2}\,\mathrm{d}x=1,$$

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin ax\cos bx}{x}\mathrm{d}x = 1;$$

再令 $b=0$，由于 $\cos bx=1$，则有

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin ax}{x}\mathrm{d}x = 1 \quad (a>0).$$

若 $a<0$，显然有

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin ax}{x}\mathrm{d}x = -\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin|a|x}{x}\mathrm{d}x = -1 \quad (a<0),$$

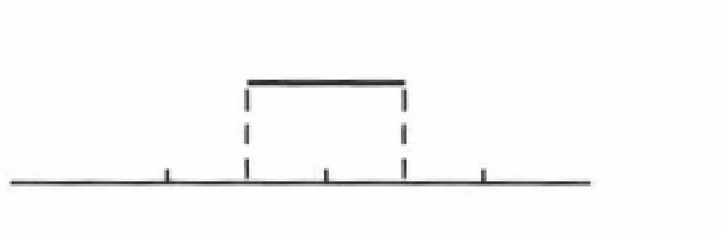
而当 $a=0$ 时，由于 $\sin ax=0$，显然有

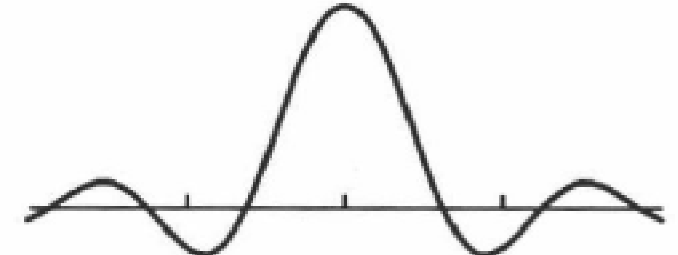
$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin ax}{x}\mathrm{d}x = 0 \quad (a=0).$$

合并写出为

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin ax}{x}\mathrm{d}x = \operatorname{sgn} a = \begin{cases}+1 & (a>0),\\ 0 & (a=0),\\ -1 & (a<0).\end{cases}$$

最后，将 $f(x)=\begin{cases}1 & (|x|<a)\\ 0 & (|x|>a)\end{cases}$ 与 $F(x)=\dfrac{2\sin ax}{x}$ 的图形绘出如下：





例 2 给定 $f(x)=\mathrm{e}^{-\frac{x^2}{2a^2}}$，求 $F(u)$.

$$F(u) = \int_{-\infty}^{\infty}\mathrm{e}^{-\frac{x^2}{2a^2}+\mathrm{i}ux}\mathrm{d}x.$$

令 $\xi=x-a^2\mathrm{i}u$，则有 $\mathrm{d}\xi=\mathrm{d}x$，以及

$$\xi^2 = x^2-2a^2\mathrm{i}ux-a^4u^2, \quad 即 \quad \frac{x^2}{2a^2}-\mathrm{i}ux = \frac{1}{2}a^2u^2+\frac{\xi^2}{2a^2};$$

故上述积分化为

$$F(u) = \mathrm{e}^{-\frac{1}{2}a^2u^2}\int_{-\infty-\mathrm{i}a^2u}^{\infty-\mathrm{i}a^2u}\mathrm{e}^{-\frac{\xi^2}{2a^2}}\mathrm{d}\xi.$$

注意到被积函数是解析的，在平行于实轴的任何带内无奇点，而且当 $|\mathrm{Re}\xi|\to\infty$ 时迅速趋于零；因而有 $\int_{-\infty-\mathrm{i}a^2u}^{\infty-\mathrm{i}a^2u}(\cdots)\mathrm{d}\xi=\int_{-\infty}^{\infty}(\cdots)\mathrm{d}\xi$（可参考 12.2.4(2) 节末之论证）.

所以，

$$F(u) = \mathrm{e}^{-\frac{1}{2}a^2u^2}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{\xi^2}{2a^2}}\mathrm{d}\xi.$$

注意到 Gauss 积分[①]

$$I=\int_{-\infty}^{\infty}\mathrm{e}^{-\gamma\xi^2}\mathrm{d}\xi=\sqrt{\frac{\pi}{\gamma}},$$

最后得到[②]

$$F(u)=\sqrt{2\pi}a\mathrm{e}^{-\frac{1}{2}a^2u^2}.$$

注意到对任意 $f(x)$ 有

$$\mathscr{F}(\alpha f(x))=\alpha F(u)=\alpha\mathscr{F}(f(x));$$

另一方面，在 Fourier 变换的定义（见 8.4.3 节）

$$F(u)=\frac{1}{C}\int_{-\infty}^{\infty}f(x)\mathrm{e}^{\mathrm{i}ux}\mathrm{d}x$$

中取不同的 C，也会引起结果的差异. 表 8.1 列出 Gauss 函数的 Fourier 变换的几个主要结果.

表 8.1 Gauss 函数的 Fourier 变换

	$f(x)$	$F(x)$	
①	$\mathrm{e}^{-x^2/2a^2}$	$\sqrt{2\pi}a\mathrm{e}^{-a^2x^2/2}$	$C=1$
②	$\frac{1}{\sqrt{2\pi}a}\mathrm{e}^{-x^2/2a^2}$	$\mathrm{e}^{-a^2x^2/2}$	$C=1$
③	$\mathrm{e}^{-x^2/2a^2}$	$a\mathrm{e}^{-a^2x^2/2}$	$C=\sqrt{2\pi}$
④	$\frac{1}{a}\mathrm{e}^{-x^2/2a^2}$	$\mathrm{e}^{-a^2x^2/2}$	$C=\sqrt{2\pi}$
⑤	$\frac{1}{\sqrt{a}}\mathrm{e}^{-x^2/2a^2}$	$\frac{1}{\sqrt{b}}\mathrm{e}^{-x^2/2b^2}$	$C=\sqrt{2\pi},\ ab=1$

应该指出：②中 $f(x)$ 称为正态分布（又称 Gauss 分布）的密度函数，其积分是归一化的. ③、④、⑤中取 $C=\sqrt{2\pi}$，具有较对称形式，⑤尤其相似. 当 $a=1$ 时，后三个的 $f(x)$ 与其 Fourier 变换 $F(x)$ 完全一样.

---

① Gauss 积分可以从

$$I^2=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{-\gamma(\xi^2+\eta^2)}\mathrm{d}\xi\mathrm{d}\eta=\int_0^{\infty}\mathrm{e}^{-\gamma\rho^2}\cdot 2\pi\rho\mathrm{d}\rho=\frac{\pi}{\gamma}$$

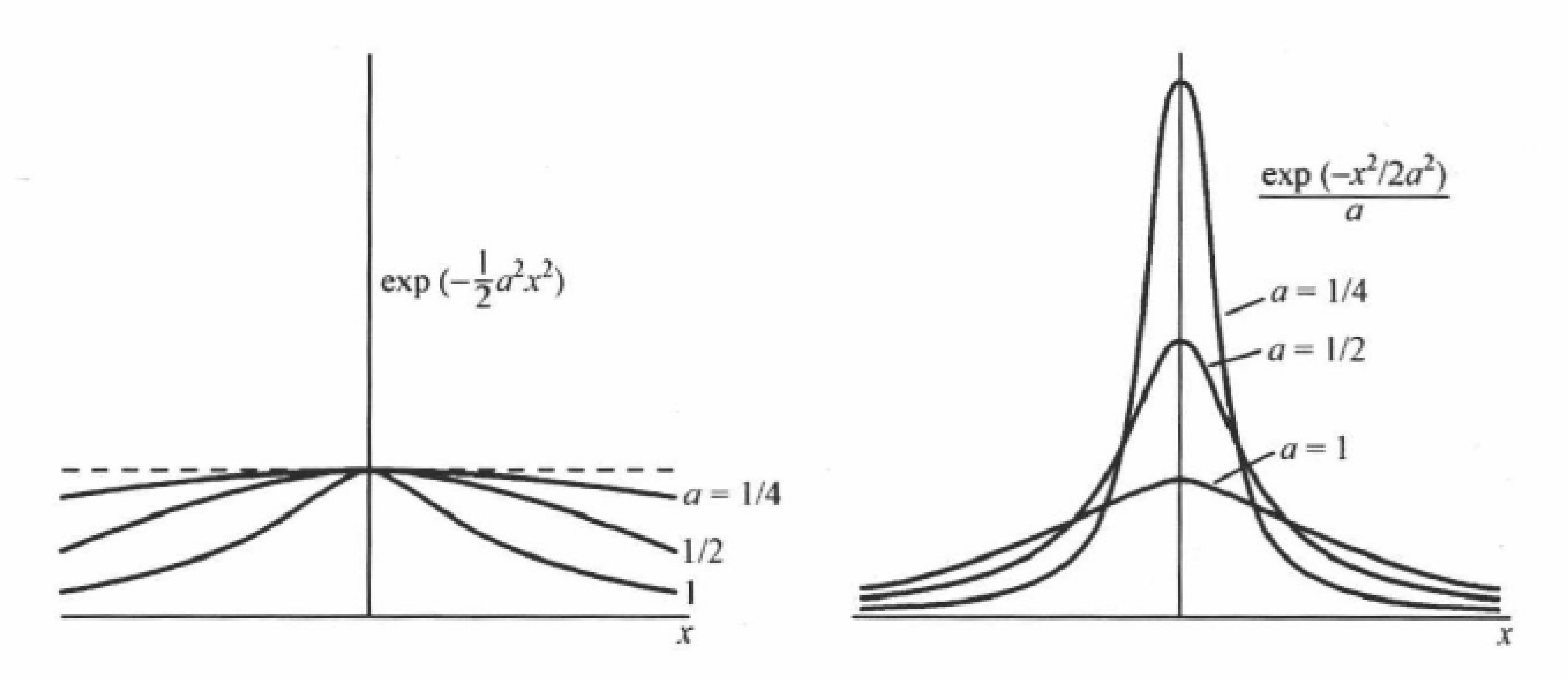
得到，求积分时引入了极坐标变换

$$\xi=\rho\cos\theta,\quad \eta=\rho\sin\theta.$$

② 由以上推导可得下列积分公式：

$$\int_{-\infty}^{\infty}\mathrm{e}^{-\gamma\xi^2+\mathrm{i}\alpha\xi}\mathrm{d}\xi=\sqrt{\frac{\pi}{\gamma}}\mathrm{e}^{-\alpha^2/4\gamma}.$$

这里给出 $a=1,1/2,1/4$ 下 $\frac{1}{a}\mathrm{e}^{-x^2/2a^2}$（下右图）和 $\mathrm{e}^{-\frac{1}{2}a^2x^2}$（下左图）的图形.



如果一个函数 $f(x)$ 及其 Fourier 变换 $F(u)$ 分别在某点 x_0 和 u_0 具有一"峰"，则二者的峰宽互成反比. 也就是说，一个"尖峰"函数具有"宽"变换，反之亦然.

关于函数的宽度可有多种定义. 若由

$$(\Delta x)^2 = \int_{-\infty}^{\infty}(x-x_0)^2\ |f(x)|^2\mathrm{d}x\Big/\int_{-\infty}^{\infty}\ |f(x)|^2\mathrm{d}x,$$

$$(\Delta u)^2 = \int_{-\infty}^{\infty}(u-u_0)^2\ |F(u)|^2\mathrm{d}u\Big/\int_{-\infty}^{\infty}\ |F(u)|^2\mathrm{d}u$$

定义其峰宽 Δx 和 Δu，则由 Schwarz 不等式（见 8.1.2 节）

$$\int |f(x)|^2\mathrm{d}x\int |g(x)|^2\mathrm{d}x \geqslant \left|\int f^*(x)g(x)\mathrm{d}x\right|^2,$$

可证明它们之间有下列重要关系：

$$\Delta x\Delta u \geqslant \frac{1}{2};$$

对于上述 Gauss 函数，可验证 $\Delta x\Delta u=1/2$.

这个重要结果 $\left(\Delta x\Delta u\geqslant\frac{1}{2}\right)$ 是量子力学中**不确定性原理**的数学基础. 例如，若令 $u=p/\hbar$，则

$$\Delta x\Delta p \geqslant \frac{1}{2}\hbar$$

是熟知的一个**不确定度关系**.

8.4.7 三维空间和四维时空的情形

物理上所处理的问题多半是在三维空间或四维时空中，现将 Fourier 变换予以推广如下：

三维空间中的 Fourier 变换是：

$$f(\boldsymbol{r}) = (2\pi)^{-3}\int F(\boldsymbol{k})\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\mathrm{d}\boldsymbol{k},$$

$$F(\boldsymbol{k}) = \int f(\boldsymbol{r})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\mathrm{d}\boldsymbol{r}.$$

四维时空中的 Fourier 变换是：

$$f(\boldsymbol{r},t) = (2\pi)^{-4}\iint F(\boldsymbol{k},\omega)\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\mathrm{d}\boldsymbol{k}\mathrm{d}\omega,$$

$$F(\boldsymbol{k},\omega) = \iint f(\boldsymbol{r},t)\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\mathrm{d}\boldsymbol{r}\mathrm{d}t.$$

8.5 运用 Fourier 分析的条件

现在来简单讨论一下运用 Fourier 分析的条件. 也就是说，在什么条件下才能将一个函数展开为 Fourier 级数或 Fourier 积分，从而可利用 Fourier 变换及其性质去解决有关问题.

前面已经指出，函数空间是矢量空间的推广，函数空间中的函数 $f(x)$ 对应于矢量空间中的矢量 $\boldsymbol{f}$. 若只考虑具有有限长度的矢量 $\boldsymbol{f}$，即 $(\boldsymbol{f},\boldsymbol{f})$ 有限；或者“长度”为有限的函数 $f(x)$，即 $\int f(x)f(x)\,\mathrm{d}x$ 有限；则这样的矢量 $\boldsymbol{f}$ 或这样的函数 $f(x)$ 属于 **Hilbert 空间**. 按照具体情况，前者属于 Euclid 空间，后者属于 $\mathbf{L}^2$ **空间（Lebesgue 平方可积函数空间**①）；即，Euclid 空间，$\mathbf{L}^2$ 空间都是 Hilbert 空间的例子. 在 Hilbert 空间中的性质，一般与以前讨论过的矢量性质很相像；所以，Hilbert 空间可视为无穷维的 Euclid 空间.

因此，$f(x)$ 可展开为 Fourier 级数或积分，运用 Fourier 分析的条件是，$f(x)$ 属于 Hilbert 空间中的 $\mathbf{L}^2$ 函数类，即 $f(x)$ 为 $\mathbf{L}^2$ 可积函数（Lebesgue 平方可积函数）；也就是说，

$$\int f(x)f(x)\,\mathrm{d}x = \text{有限}.$$

M. Plancherel 确立了下面定理：

$$\int_{-\infty}^{\infty} |f(x)|^2\mathrm{d}x = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(u)|^2\mathrm{d}u$$

成立的条件是两个积分存在，后来 T. Carleman 又证明了只要其中一个积分存在

---

① 对有界函数 $f(x)$，其 **Lebesgue 积分**定义为

$$\int f(x)\,\mathrm{d}x = \lim_{\lambda\to 0}\sum_{i=1}^{n}\xi_i\mu(\{x\,|\,\xi_{i-1}\leqslant f(x)<\xi_i\}),$$

其中 $\lambda=\max\limits_{1\leqslant i\leqslant n}(\xi_i-\xi_{i-1})$，而 $\mu(\{x|\cdots\})$ 为满足后面条件的“点集”的“长度”，称为 **Lebesgue 测度**.

即可确立. 所以,

$$\int f^*(x)f(x)\,\mathrm{d}x = \text{有限}$$

的条件常称为 $\mathbf{L}^2$ 函数类的 Plancherel 理论,而上述定理又称为 Plancherel 定理.

属于 Hilbert 空间($\mathbf{L}^2$ 空间)的函数的例子有:

(i) 量子力学中的波函数 $\psi(x)$,要求

$$\int \psi^*(x)\psi(x)\mathrm{d}x = \text{有限},$$

通常归一化为 1.

(ii) 电场 $\boldsymbol{E}(\boldsymbol{x})$,因为要求总电能为有限,即

$$\int \frac{1}{2}\varepsilon E^2\,\mathrm{d}x = \text{有限}.$$

(iii) 凡是具有振幅性质的物理量,都可运用 Fourier 分析来处理,因为

$$\int(\text{振幅})^2\,\mathrm{d}x \propto \text{能量} \neq \infty.$$

另一方面,由 Fourier 变换公式:

$$\begin{cases} F(u) = \displaystyle\int_{-\infty}^{\infty} \underbrace{f(x)}_{\text{给①}}\mathrm{e}^{\mathrm{i}ux}\,\mathrm{d}x, \\ f(x) = \underbrace{\displaystyle\frac{1}{2\pi}\int_{-\infty}^{\infty} F(u)\,\mathrm{e}^{-\mathrm{i}ux}\,\mathrm{d}u}_{\text{算②}}; \end{cases}$$

首先将所给的 $f(x)$用 Fourier 变换公式算出 $F(u)$,再将此 $F(u)$用 Fourier 逆变换公式反算出 $\bar{f}(x)$(记为 $\bar{f}(x)$,以示区别),即

$$\bar{f}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(u)\,\mathrm{e}^{-\mathrm{i}ux}\,\mathrm{d}u = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(x')\,\mathrm{e}^{\mathrm{i}ux'}\,\mathrm{d}x'\right\}\mathrm{e}^{-\mathrm{i}ux}\,\mathrm{d}u;$$

若 $\int f(x)f(x)\,\mathrm{d}x = \text{有限}$, 则可证明①

$$\int\{\underset{\text{给①}}{f(x)} - \underset{\text{算②}}{\bar{f}(x)}\}^2\,\mathrm{d}x = 0;$$

即函数空间中 $f(x)$与 $\bar{f}(x)$之距离为零. 通常称 Fourier 积分之 Plancherel 理论为 $\mathbf{L}^2$ 理论,或用 l. i. m. 表示平均收敛,

$$\text{l. i. m. } \bar{f}(x) = f(x);$$

严格地说来应为均方收敛.

这里牵涉到极限的概念. 给定 $f(x)$使 $F(u)$积分存在,但从 $F(u)$ 对 u 积分算

---

① E. C. Titchmarsh, *Theory of Fourier Integrals*, Clarendon Press, Oxford, England, 1937, §3.3.

$f(x)$时，对 u 的无穷积分区间

$$\int_{-\infty}^{\infty} F(u)\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}u \quad 像 \quad \int_{-b}^{b} F(u)\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}u$$

的 $b\to\infty$极限（下限也可以独立进行，用 $a\to-\infty$）. 令

$$\tilde{f}(x,b) \equiv \frac{1}{2\pi}\int_{-b}^{b}\left\{\int_{-\infty}^{\infty} f(x')\mathrm{e}^{\mathrm{i}ux'}\mathrm{d}x'\right\}\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}u,$$

则有

$$\underset{b\to\infty}{\mathrm{l.i.m.}}\ \tilde{f}(x,b) = f(x);$$

即前面应为→0. 粗糙地用

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}u(x'-x)}\mathrm{d}u = \delta(x'-x) \quad 及$$

$$\int_{-\infty}^{\infty}\delta(x'-x)f(x')\mathrm{d}x' = f(x)$$

表示.

8.6　相关积分变换

下面介绍两个与 Fourier 变换相关的积分变换.

8.6.1　Laplace 变换

对于 Fourier 变换

$$G(u) = \int_{-\infty}^{\infty} g(x)\mathrm{e}^{\mathrm{i}ux}\mathrm{d}x,$$

若限制函数 $g(x)$使得

$$g(x) = \begin{cases} f(x), & x>0, \\ 0, & x<0; \end{cases}$$

并令 $x\to t$，$\mathrm{i}u\to -p$，则得到

$$F_{\mathrm{L}}(p) = \int_{0}^{\infty} f(t)\mathrm{e}^{-pt}\mathrm{d}t,$$

此式称为函数 $f(t)$的 **Laplace 变换**；或记作

$$\mathscr{L}\{f(t)\} = F_{\mathrm{L}}(p).$$

同时，对于 Fourier 逆变换

$$g(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(u)\mathrm{e}^{-\mathrm{i}ux}\mathrm{d}u,$$

则相应变为 **Laplace 逆变换**

$$f(t) = \frac{1}{2\pi\mathrm{i}}\int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} F_{\mathrm{L}}(p)\mathrm{e}^{pt}\mathrm{d}p, \quad t>0;$$

其中积分路径沿虚轴.

另一方面,若 $f(t)$局部可积,$e^{-\gamma t}|f(t)|$在区间$(0,\infty)$上对某个 $\gamma=c>0$ 可积,则

$$F_{\mathrm{L}}(p)=\int_0^\infty f(t)\mathrm{e}^{-pt}\mathrm{d}t,\quad \mathrm{Re}p\geqslant c>0,$$

$$f(t)=\frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty}F_{\mathrm{L}}(p)\mathrm{e}^{pt}\mathrm{d}p,\quad t>0,$$

其中积分路径沿 $\mathrm{Re}p=c>0$ 的直线.

下面举几个例子:

① $\mathscr{L}\{\mathrm{e}^{at}\}=\int_0^\infty \mathrm{e}^{at}\mathrm{e}^{-pt}\mathrm{d}t=\frac{1}{p-a},\mathrm{Re}p>\mathrm{Re}a.$

由于当 $a=0$ 时,$\mathrm{e}^{at}=1$,所以

$$\mathscr{L}\{1\}=\frac{1}{p},\quad \mathrm{Re}p>0.$$

② $\mathscr{L}\{\mathrm{e}^{\mathrm{i}bt}\}=\int_0^\infty \mathrm{e}^{\mathrm{i}bt}\mathrm{e}^{-pt}\mathrm{d}t=\frac{1}{p-\mathrm{i}b}=\frac{p+\mathrm{i}b}{p^2+b^2},\quad \mathrm{Re}p>|\mathrm{Im}b|.$

由于 $\mathrm{e}^{\mathrm{i}bt}=\cos bt+\mathrm{i}\sin bt$,所以

$$\mathscr{L}\{\cos bt\}=\frac{p}{p^2+b^2},\quad \mathrm{Re}p>|\mathrm{Im}b|,$$

$$\mathscr{L}\{\sin bt\}=\frac{b}{p^2+b^2},\quad \mathrm{Re}p>|\mathrm{Im}b|.$$

③ $\mathscr{L}\{t^\alpha\}=\int_0^\infty t^\alpha\mathrm{e}^{-pt}\mathrm{d}t=\frac{1}{p^{\alpha+1}}\int_0^\infty x^\alpha\mathrm{e}^{-x}\mathrm{d}x=\frac{\Gamma(\alpha+1)}{p^{\alpha+1}},\quad \mathrm{Re}p>0.$

④ 另外,对 Laplace 变换

$$F_{\mathrm{L}}(p)=\int_0^\infty f(t)\mathrm{e}^{-pt}\mathrm{d}t,\quad \mathrm{Re}p>0$$

两边取导数,容易得到

$$F_{\mathrm{L}}^{(n)}(p)=\left(\frac{\partial}{\partial p}\right)^n\int_0^\infty f(t)\mathrm{e}^{-pt}\mathrm{d}t=(-)^n\int_0^\infty t^nf(t)\mathrm{e}^{-pt}\mathrm{d}t,\quad \mathrm{Re}p>0,$$

或者,

$$\mathscr{L}\{t^nf(t)\}=(-)^nF_{\mathrm{L}}^{(n)}(p),\quad \mathrm{Re}p>0.$$

8.6.2 Mellin 变换

对于 Fourier 变换:

$$G(u)=\int_{-\infty}^\infty g(x)\mathrm{e}^{\mathrm{i}ux}\mathrm{d}x,$$

若令 $\mathrm{e}^x\to y,\mathrm{i}u\to s$,并且 $g(x)\to f(y)$,于是有

$$F_M(s)=\int_0^\infty f(y)y^{s-1}\mathrm{d}y,\quad \sigma_0<\mathrm{Re}s<\sigma_1,$$

后面的条件表示积分的绝对收敛区域；此式称为函数 $f(y)$ 的 **Mellin 变换**，或记作

$$\mathscr{M}\{f(y)\}=F_M(s).$$

同时，由 Fourier 逆变换可推得相应的 Mellin 逆变换：

$$f(y)=\frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty}F_M(s)y^{-s}\mathrm{d}s,\quad \sigma_0<c<\sigma_1.$$

下面举几个简单例子：

① $\mathscr{M}\{\mathrm{e}^{-ay}\}=\dfrac{\Gamma(s)}{a^s},\quad \mathrm{Re}\,a>0,\mathrm{Re}\,s>0.$

② $\mathscr{M}\{\cos ay\}=\dfrac{\Gamma(s)}{a^s}\cos\left(\dfrac{\pi s}{2}\right),\quad a>0,\quad 0<\mathrm{Re}\,s<1;$

$\mathscr{M}\{\sin ay\}=\dfrac{\Gamma(s)}{a^s}\sin\left(\dfrac{\pi s}{2}\right),\quad a>0,\quad -1<\mathrm{Re}\,s<1.$

③ $\mathscr{M}\left\{\dfrac{1}{1+y}\right\}=\dfrac{\pi}{\sin(\pi s)},\quad 0<\mathrm{Re}\,s<1;$

$\mathscr{M}\left\{\dfrac{1}{1-y}\right\}=\dfrac{\pi\cos(\pi s)}{\sin(\pi s)},\quad 0<\mathrm{Re}\,s<1.$

④ $\mathscr{M}\{\ln(1+y)\}=\dfrac{\pi}{s\sin(\pi s)},\quad -1<\mathrm{Re}\,s<0;$

$\mathscr{M}\{\ln|1-y|\}=\dfrac{\pi\cos(\pi s)}{s\sin(\pi s)},\quad -1<\mathrm{Re}\,s<0.$

习　题

8.1 求乘积 $f(x)g(x)$ 的 Fourier 变换.

解：

$$\begin{aligned}\mathscr{F}(f(x)g(x))&=\int_{-\infty}^{\infty}f(x)g(x)\mathrm{e}^{\mathrm{i}ux}\mathrm{d}x\\&=\int_{-\infty}^{\infty}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}F(u')\mathrm{e}^{-\mathrm{i}u'x}\mathrm{d}u'\right]\cdot\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}G(u'')\mathrm{e}^{-\mathrm{i}u''x}\mathrm{d}u''\right]\mathrm{e}^{\mathrm{i}ux}\mathrm{d}x\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[F(u')G(u'')\mathrm{d}u'\mathrm{d}u'']\cdot\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(u-u'-u'')x}\mathrm{d}x\right]\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(u')G(u'')\delta(u-u'-u'')\mathrm{d}u'\mathrm{d}u''\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(u')G(u-u')\mathrm{d}u'\\&=\frac{1}{2\pi}\{F*G\}(u).\end{aligned}$$

8.2 求(1) $\frac{\mathrm{d}^n f(x)}{\mathrm{d}x^n}$ 和 (2) $(\mathrm{i}x)^n f(x)$的 Fourier 变换.

解 (1) $\mathscr{F}\left(\frac{\mathrm{d}^n f(x)}{\mathrm{d}x^n}\right)=(-\mathrm{i}u)^n F(u)$,

(2) $\mathscr{F}((\mathrm{i}x)^n f(x))=\frac{\mathrm{d}^n F(u)}{\mathrm{d}u^n}$.

8.3 求 $\sum_{n=-\infty}^{\infty}\delta(x-nx_0)$ 的 Fourier 级数及 Fourier 变换①.

解 这是以等距 x_0 分布的 δ 函数. 求其 Fourier 级数可取区间$\left(-\frac{1}{2}x_0,\frac{1}{2}x_0\right)$作展开,这时三角函数系中 x 应变为$\frac{2\pi}{x_0}x$. 于是

$$a_m=\frac{2}{x_0}\int_{-\frac{1}{2}x_0}^{\frac{1}{2}x_0}\delta(x)\cos m\frac{2\pi}{x_0}x\mathrm{d}x=\frac{2}{x_0},\quad b_n=0;$$

所以,

$$\sum_{n=-\infty}^{\infty}\delta(x-nx_0)=\frac{2}{x_0}\left\{\frac{1}{2}+\sum_{m=1}^{\infty}\cos m\frac{2\pi}{x_0}x\right\}=\frac{1}{x_0}\sum_{m=-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}m\frac{2\pi}{x_0}x};$$

最后一步利用了公式 $\cos\theta=\frac{1}{2}(\mathrm{e}^{\mathrm{i}\theta}+\mathrm{e}^{-\mathrm{i}\theta})$. 其实,直接采用 Fourier 级数的复数形式可更方便地得到后一结果.

现在再来利用上述结果求它的 Fourier 变换.

$$\begin{aligned}\mathscr{F}\left(\sum_{n=-\infty}^{\infty}\delta(x-nx_0)\right)&=\int_{-\infty}^{\infty}\left[\frac{1}{x_0}\sum_{m=-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}m\frac{2\pi}{x_0}x}\right]\mathrm{e}^{\mathrm{i}ux}\mathrm{d}x\\&=\frac{2\pi}{x_0}\sum_{m=-\infty}^{\infty}\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}\left[u-m\frac{2\pi}{x_0}\right]x}\mathrm{d}x\right\}\\&=\frac{2\pi}{x_0}\sum_{m=-\infty}^{\infty}\delta\left(u-m\frac{2\pi}{x_0}\right).\end{aligned}$$

8.4 求 $\sum_{n=-\infty}^{\infty}\delta(x-nx_0)f(x)$ 的 Fourier 变换.

解 利用习题 8.1 和 8.3 的结果,容易得到

$$\begin{aligned}&\mathscr{F}\left(\sum_{n=-\infty}^{\infty}\delta(x-nx_0)f(x)\right)\\&\quad=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left\{\left[\frac{2\pi}{x_0}\sum_{m=-\infty}^{\infty}\delta\left(u'-m\frac{2\pi}{x_0}\right)\right]F(u-u')\right\}\mathrm{d}u'\end{aligned}$$

① 习题 8.3 和 8.4 可参考: D. C. Champeney, *Fourier Transforms and their Physical Applications*, Academic Press, 1973; pp. 7,19,32,218.

$$= \frac{1}{x_0}\sum_{n=-\infty}^{\infty} F\left(u - m\frac{2\pi}{x_0}\right).$$

8.5 证明下列 Poisson 求和公式：

$$\sum_{n=-\infty}^{\infty} f(\alpha n) = \frac{1}{\alpha}\sum_{m=-\infty}^{\infty} F\left(m\frac{2\pi}{\alpha}\right).$$

证 令上题中 $x_0=\alpha$，求其逆变换，并在两边对 x 求积分即可得证.

8.6 Laplace 变换的卷积定理：

$$\mathscr{L}\left\{\int_0^t f(\tau)g(t-\tau)\mathrm{d}\tau\right\} = \mathscr{L}\{f(t)\}\cdot\mathscr{L}\{g(t)\}.$$

8.7 求 $f^{(n)}(t)=\frac{\mathrm{d}^n}{\mathrm{d}t^n}f(t)$ 的 Laplace 变换.

解

$$\begin{aligned}\mathscr{L}\{f'(t)\} &= \int_0^{\infty} f'(t)\mathrm{e}^{-pt}\mathrm{d}t \\ &= [f(t)\mathrm{e}^{-pt}]_0^{\infty} + p\int_0^{\infty} f(t)\mathrm{e}^{-pt}\mathrm{d}t \\ &= p\mathscr{L}\{f(t)\} - f(0).\end{aligned}$$

类似地可得

$$\mathscr{L}\{f''(t)\} = p\mathscr{L}\{f'(t)\} - f'(0),$$
$$\vdots$$
$$\mathscr{L}\{f^{(n)}(t)\} = p\mathscr{L}\{f^{(n-1)}(t)\} - f^{(n-1)}(0);$$

于是，容易求得

$$\mathscr{L}\{f^{(n)}(t)\} = p^n\mathscr{L}\{f(t)\} - \sum_{m=1}^{n} p^{m-1}f^{(n-m)}(0).$$

8.8 试用 Mellin 逆变换公式验证：

(1) $\mathscr{M}\left\{\frac{1}{1+y}\right\}=\frac{\pi}{\sin(\pi s)}$，　(2) $\mathscr{M}\left\{\frac{1}{1-y}\right\}=\frac{\pi\cos(\pi s)}{\sin(\pi s)}$.

证 若令 $\theta=\pm1,\theta'=\frac{1}{2}(\theta+1)$，则可将上列二式合并写成

$$\mathscr{M}\left\{\frac{1}{1-\theta y}\right\} = \frac{\pi\cos^{\theta'}(\pi s)}{\sin(\pi s)}.$$

现在由 $F_{\mathrm{M}}(s)=\frac{\pi\cos^{\theta'}(\pi s)}{\sin(\pi s)}$ 利用 Mellin 逆变换公式来求 $f(y)$：

$$f(y) = \frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty}\frac{\pi\cos^{\theta'}(\pi s)}{\sin(\pi s)}y^{-s}\mathrm{d}s \quad (0<c<1).$$

被积函数在实轴上有来自 $\sin(\pi s)$ 的简单极点 $s=\pm l$，其留数 $r_{\pm l}$ 为：

$$\begin{aligned}r_{\pm l} &= \lim_{s=\pm l}(s\mp l)\frac{\pi\cos^{\theta'}(\pi s)}{\sin(\pi s)}y^{-s} \\ &= \cos^{(\theta'-1)}(\pm\pi l)y^{\mp l}\end{aligned}$$

$$= (-)^{l(\theta'-1)} y^{\mp l} = (\theta y)^{\mp l};$$

最后一步是因为 $\theta' = \frac{1}{2}(\theta+1)$ 而当 $\theta = \pm 1$ 时有 $(-)^{\theta'-1} = \theta$.

当 $y<0$ 时,采用向左闭合的周线,得到

$$f(y) = \frac{1}{2\pi i} 2\pi i \sum_{l=0}^{\infty} r_{-l} = \sum_{l=0}^{\infty} (\theta y)^l = \frac{1}{1-\theta y}.$$

当 $y>1$ 时,采用向右闭合的周线,得到

$$f(y) = -\sum_{l=1}^{\infty} r_{+l} = -\sum_{l=1}^{\infty} (\theta y)^{-l}$$

$$= 1 - \sum_{l=0}^{\infty} (\theta y)^{-l} = 1 - \frac{1}{1-(\theta y)^{-1}} = \frac{1}{1-\theta y};$$

求和号前的"$-$"号是由于周线为顺时针方向. 得证.

8.9 验证下列 Mellin 变换:

(1) $\mathscr{M}\{\ln(1+y)\} = \frac{\pi}{s\sin(\pi s)}$ $(-1<\mathrm{Re}\, s<0)$,

(2) $\mathscr{M}\{\ln|1-y|\} = \frac{\pi\cos(\pi s)}{s\sin(\pi s)}$ $(-1<\mathrm{Re}\, s<0)$.

为了验证此结果,先推导一个公式:

$$\mathscr{M}\left\{\int_0^y f(x)\mathrm{d}x\right\}(s) = \int_0^{\infty}\left[\int_0^y f(x)\mathrm{d}x\right] y^{s-1}\mathrm{d}y$$

$$= \left[\frac{y^s}{s}\int_0^y f(x)\mathrm{d}x\right]_0^{\infty} - \frac{1}{s}\int_0^{\infty} f(y) y^s \mathrm{d}y$$

$$= -\frac{1}{s}\mathscr{M}\{f(y)\}(s+1),$$

为了使积分项为零,要求 $\mathrm{Re}\, s<0$. 于是,由于

$$\ln(1-\theta y) = -\theta\int_0^y \frac{\mathrm{d}x}{1-\theta x},$$

因而,

$$\mathscr{M}\{\ln(1-\theta y)\}(s) = \frac{\theta}{s}\mathscr{M}\left\{\frac{1}{1-\theta y}\right\}(s+1)$$

$$= \frac{\theta}{s}\frac{\pi\cos^{\theta'}(\pi+\pi s)}{\sin(\pi+\pi s)} \quad (0<\mathrm{Re}(s+1)<1)$$

$$= (-)^{\theta'-1}\theta\frac{\pi}{s}\frac{\cos^{\theta'}(\pi s)}{\sin(\pi s)} = \frac{\pi}{s}\frac{\cos^{\theta'}(\pi s)}{\sin(\pi s)} \quad (-1<\mathrm{Re}\, s<0).$$

这里用了 $(-)^{\theta'-1} = \theta$, $\theta^2 = 1$. 用 $\theta = \pm$ 和 $\theta' = \frac{1}{2}(\theta+1)$ 代入后即得所要验证之结果.

第 9 章　变　分　法

变分法的最基本概念是泛函，后者是函数概念的发展. 函数有极值问题，而正是对泛函极值问题的研究导致变分法的建立.

9.1　泛　　函

9.1.1　定义和例子

函数以数为自变元来定义，泛函则以函数为自变元来定义.

对于函数 $f(x)$[例如，$f(x)=x^2$]，当数 x 在区间$[a,b]$上一点之值给定时，即得 $f(x)$之对应值，或者说，函数 $f(x)$是点函数；当 x 在$[a,b]$上变动时，$f(x)$值相应地变动. 但是，对于泛函 $f[x(t)]$[例如，$f[x(t)]=\int_0^1\{x(t)\}^2\mathrm{d}t$]，只有当函数 $x(t)$在数 t 的变化区间（上例中为$[0,1]$）上各点之值全给定时，才能够得到泛函 $f[x(t)]$ 的一个对应值，或者说，泛函 $f[x(t)]$是线函数；而只有当 $x(t)$之函数形式变动时，泛函 $f[x(t)]$之值才会有所变动；故称为 t 之函数 $x(t)$的**泛函**. 这里还要指出，泛函与复合函数之间的区别. 对于复合函数 $f(x(t))$[例如，$f(x(t))=\{x(t)\}^2$]，它所表示的是：x 是 t 的函数，f 又是 x 的函数，即，复合函数 $f(x(t))$是函数 $x(t)$的函数；然而，当给定 $t=t_0\,(a\leqslant t_0\leqslant b)$时，函数 $x(t)$即得一值 $x(t_0)$，同时给出复合函数 $f(x(t))$的值 $f(x(t_0))$；也就是说，复合函数仍是点函数，并非泛函那样的线函数. 为了在记号上显示这种区别，将泛函记作 $f[x]$或 $f[x(t)]$（用方括号），以区别于函数 $f(x)$ 或复合函数 $f(x(t))$（用圆括号）.

下面举出泛函的若干例子.

(i)
$$f[x(t)]=\int_0^1\{x(t)\}^2\mathrm{d}t,$$

它与 $f=\sum_{i=1}^{n}x_i^2$ 相像.

(ii)
$$g(x)=\int_a^b K(x,y)h(y)\mathrm{d}y,$$

它具有第一类 Fredholm 积分方程或积分变换的形式，这里则将 $g(x)$ 看做 $h(y)$的泛函以及 x 的函数，可记作 $g[h(y)](x)$.

(iii) $$K[\varphi(x)]=\int_a^b\int_a^b K(x,y)\varphi(x)\varphi(y)\mathrm{d}x\mathrm{d}y,$$

它与 $\sum_{i=1}^{n}\sum_{j=1}^{n}K_{ij}\varphi_i\varphi_j$ 相像，是 $\varphi(x)$的泛函.

(iv) $$I[z(x,y)]=\iint_R\sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}\mathrm{d}x\mathrm{d}y,$$

它代表曲面积，是二元函数 $z(x,y)$的泛函.

(v) $$K[\varphi(x,y),\psi(y)]=\int_a^b\int_a^b K(x,y)\varphi(x,y)\psi(y)\mathrm{d}x\mathrm{d}y,$$

$$K[\varphi,\psi]=\iint K(\boldsymbol{x},y)\varphi(\boldsymbol{x})\psi(y)\mathrm{d}\boldsymbol{x}\mathrm{d}y,$$

它们是两个函数 φ,ψ 的泛函.

(vi) 泛函 $f[x(t)]$的一般形式为

$$f[x(t)]=\int_a^b F(x(t),\dot{x}(t),\cdots)\mathrm{d}t,$$

$$f[x(t)]=\int_a^b\cdots\int_a^b F(x(t'),\cdots,x(t^{(n)});\ \dot{x}(t'),\cdots,\dot{x}(t^{(n)});\cdots)\mathrm{d}t'\cdots\mathrm{d}t^{(n)},$$

函数 F 中后面的…可包含 x 的高阶导数等.

应该指出，凡是具有历史性的现象，如滞后效应、金属性质等物理现象，以及国家经济状况、人口状况等社会现象，都要对过去历史进行积分才能知道它们现在的情况；也就是说，需用泛函来描述，这就是泛函的意义.

9.1.2 无穷个变量的函数

考虑泛函 $f[x(t)]$的下列一般形式

$$f[x(t)]=\int_a^b F(x(t),\dot{x}(t),\cdots)\mathrm{d}t.$$

若令 $\Delta t=(b-a)/N$，$t_j=a+j\Delta t(t_0=a,t_N=b)$，以及令

$$x(t_i)=x_i\quad(i=1,2,\cdots,N),$$

$$\dot{x}(t_i)=(x_{i+1}-x_i)/\Delta t,$$

$$\ddot{x}(t_i)=(x_{i+2}-2x_{i+1}+x_i)/(\Delta t)^2,$$

等等；于是，按照 Riemann 积分的概念，

$$\begin{aligned}f[x(t)]&=\int_a^b F(x(t),\dot{x}(t),\cdots)\mathrm{d}t\\&=\lim_{N\to\infty}\sum_{i=1}^{N}F(x(t_i),\dot{x}(t_i),\cdots)\Delta t\\&=\lim_{N\to\infty}\sum_{i=1}^{N}F(x_i,(x_{i+1}-x_i)/\Delta t,\cdots)\Delta t\end{aligned}$$

$$= \lim_{N\to\infty}\mathscr{F}(x_1,\cdots,x_N);$$

也就是说，一个泛函可以看做是具有无穷个变量的一个函数.

9.1.3 无穷维函数空间中的函数

另一方面，泛函与函数的关系就好像函数空间与 N 维空间的关系一样.

例如，考虑泛函

$$f[x(t)] = \int_{-\pi}^{\pi}\{x(t)\}^2\mathrm{d}t,$$

若将函数 $x(t)$ 用三角函数系表示为

$$x(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n\cos nt + b_n\sin nt),$$

则泛函 $f[x(t)]$ 可表示为

$$f[x(t)] = \pi\left\{\frac{1}{2}a_0^2 + \sum_{n=1}^{\infty}(a_n^2 + b_n^2)\right\}.$$

正交归一函数系(例中为三角函数系)构成(无穷维)函数空间的"坐标系"，一个函数(如 $x(t)$)是在函数空间的一个"点"(由"坐标"a_0,a_n,b_n 表示)，而一个泛函(如例中的 $f[x(t)]$)则是此函数空间中的"点函数"；或者说，一个泛函就是无穷个变量(例中为 a_0,a_n,b_n)的函数.

9.2 变分法的意义

变分法的意义就是求泛函的极值问题.

9.2.1 函数的极值问题

对于一个普通多元函数 $f(x_1,x_2,\cdots,x_n)$，使其一阶导数均为零的点：

$$\frac{\partial f}{\partial x_1} = 0,\quad \frac{\partial f}{\partial x_2} = 0,\quad \cdots,\quad \frac{\partial f}{\partial x_n} = 0;$$

或者，写成矢量的形式

$$\begin{pmatrix}\dfrac{\partial f}{\partial x_1}\\ \dfrac{\partial f}{\partial x_2}\\ \vdots\\ \dfrac{\partial f}{\partial x_n}\end{pmatrix} = 0,\quad 即\quad \mathrm{grad}f = 0;$$

称为**平稳点**；由上述条件解出，可有若干个. 设有某一平稳点 $x^{(0)}$ 使其二阶导数均

不为零(或更一般地,直至某偶阶导数才不为零),则 $f(x^{(0)})$是在该点邻域的极值,偶阶导数均<0时为极大值而>0时则为极小值;这样的点称为**极值点**. 当然,还有其他的情况.

下面举几个简单例子.

(i) $f(x_1,\cdots,x_n)=\sum_{i=1}^{n}x_i^2$ 在 $x_1=x_2=\cdots=x_n=0$ 的点有极小值.

(ii) $f(x)=x^4-x^2$ 有三个平稳点,$x=0,\pm\frac{1}{2}\sqrt{2}$;由于 $f''(0)=-2<0$, $f''\left(\pm\frac{1}{2}\sqrt{2}\right)=4>0$,故 $f(x)$在 $x=0$ 邻域是极大值,而在 $x=\pm\frac{1}{2}\sqrt{2}$邻域是极小值.

(iii) $f(x)=x^3$,在 $x=0$ 时有 $f'(0)=f''(0)=0$,而奇阶导数 $f'''(0)=6\neq0$,它是函数 $f(x)$的拐点.

(iv) $f(x_1,x_2)=x_1^2-x_2^2$ 有一个平稳点 $x_1=x_2=0$,但由于$\frac{\partial^2 f}{\partial x_1^2}=2>0$ 而$\frac{\partial^2 f}{\partial x_2^2}=-2<0$,它是函数 $f(x_1,x_2)$的鞍点.

9.2.2 泛函的极值问题

前面曾指出,函数可看做函数空间的"点",泛函则可看做此函数空间中的"点函数";或者

泛函[(一个参量的)变函数].

因此,要求泛函的平稳值,先要规定变函数在函数空间中的变化区域,然后规定如何来求泛函的导数,以及确定泛函之平稳值的条件. 这样的方法就称为**变分法**.

9.3 Euler 变分方程

9.3.1 泛函的变分导数

对于普通函数$f(x)$,当 x 在 x_0 附近变化时,可作 Taylor 展开:

$$f(x_0+\varepsilon)=f(x_0)+\varepsilon\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)_{x=x_0}+\frac{\varepsilon^2}{2}\left(\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right)_{x=x_0}+\cdots,$$

而$\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)_{x=x_0}=0$ 是 $x=x_0$ 为平稳点的条件. 或者,一般地,当数 x 作 $x\to x+\delta x$ 变动时,$f(x)$变为

$$f(x+\delta x)=f(x)+\delta x\frac{\mathrm{d}f}{\mathrm{d}x}+\frac{(\delta x)^2}{2!}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}+\cdots,$$

其一阶变分是

$$\delta f = \delta x \frac{\mathrm{d}f}{\mathrm{d}x}.$$

对于多元函数 $f(x_1, x_2, \cdots)$，类似地有

$$f(x_1+\delta x_1, x_2+\delta x_2, \cdots) = f(x_1, x_2, \cdots) + \left(\delta x_1 \frac{\partial f}{\partial x_1} + \delta x_2 \frac{\partial f}{\partial x_2} + \cdots\right) + \cdots,$$

其一阶变分是

$$\delta f = \delta \boldsymbol{x} \cdot \mathrm{grad} f.$$

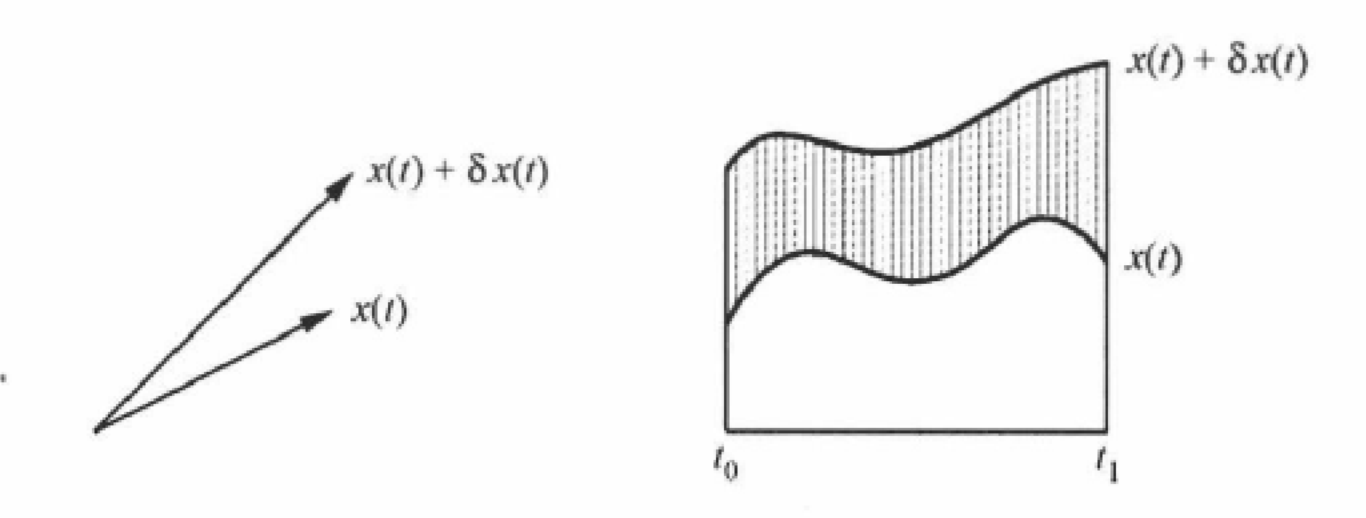
因此，函数 $f(\boldsymbol{x})$的平稳点条件由其一阶变分为零给出，即

$$\delta f = 0 \rightarrow \mathrm{grad} f = 0.$$

现在来讨论泛函

$$f[x(t)] = \int_{t_0}^{t_1} F(x, \dot{x}, t)\mathrm{d}t.$$

在函数空间，同样可以考虑其中一点 $x(t)$之变化：$x(t) \rightarrow x(t) + \delta x(t)$如下图，



于是，可类似地写出**泛函之一阶变分**为

$$\delta f = \int_{t_0}^{t_1} \delta x(t)[F]_{x(t)}\mathrm{d}t,$$

其中$[F]_{x(t)}$称为**变分导数**或**泛函导数**，它本身代表函数空间中的梯度. 同时，泛函 $f[x(t)]$的平稳点条件由其一阶变分为零给出，即

$$\delta f = 0 \rightarrow [F]_{x(t)} = 0,$$

也就是说，要求其变分导数为零；由此结果得出的方程称为 **Euler 方程**. 应该指出，以上是就**固定边界**而言，对于**自由边界**或**变端点**的情况，会有相应修正，下面分别予以讨论.

9.3.2 Euler 变分方程 · 边界条件

现在来具体讨论下列泛函：

$$f[x(t)] = \int_{t_0}^{t_1} F(x, \dot{x}, t)\mathrm{d}t.$$

当 $x(t) \rightarrow x(t) + \delta x(t)$时，$f[x(t)]$变为

$$f[x(t)+\delta x(t)]=\int_{t_0}^{t_1}F\left(x+\delta x,\frac{\mathrm{d}x}{\mathrm{d}t}+\frac{\mathrm{d}(\delta x)}{\mathrm{d}t},t\right)\mathrm{d}t$$

$$=\int_{t_0}^{t_1}\left\{F\left(x,\frac{\mathrm{d}x}{\mathrm{d}t},t\right)+\left[\frac{\partial F}{\partial x}\delta x+\frac{\partial F}{\partial\frac{\mathrm{d}x}{\mathrm{d}t}}\cdot\frac{\mathrm{d}(\delta x)}{\mathrm{d}t}\right]+\cdots\right\}\mathrm{d}t,$$

所以,泛函 $f[x(t)]$的一阶变分是

$$\delta f[x]=\int_{t_0}^{t_1}\left[\frac{\partial F}{\partial x}\delta x+\frac{\partial F}{\partial\frac{\mathrm{d}x}{\mathrm{d}t}}\frac{\mathrm{d}(\delta x)}{\mathrm{d}t}\right]\mathrm{d}t$$

$$=\int_{t_0}^{t_1}\delta x\left[\frac{\partial F}{\partial x}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial\frac{\mathrm{d}x}{\mathrm{d}t}}\right]\mathrm{d}t+\left\{\delta x\frac{\partial F}{\partial\frac{\mathrm{d}x}{\mathrm{d}t}}\right\}\Bigg|_{t=t_0}^{t=t_1}$$

$$=\int_{t_0}^{t_1}\delta x(t)[F]_{x(t)}\mathrm{d}t+\delta x(t_1)\frac{\partial F}{\partial\dot{x}}\bigg|_{t=t_1}-\delta x(t_0)\frac{\partial F}{\partial\dot{x}}\bigg|_{t=t_0},$$

其中$[F]_{x(t)}$为泛函的**变分导数**,即

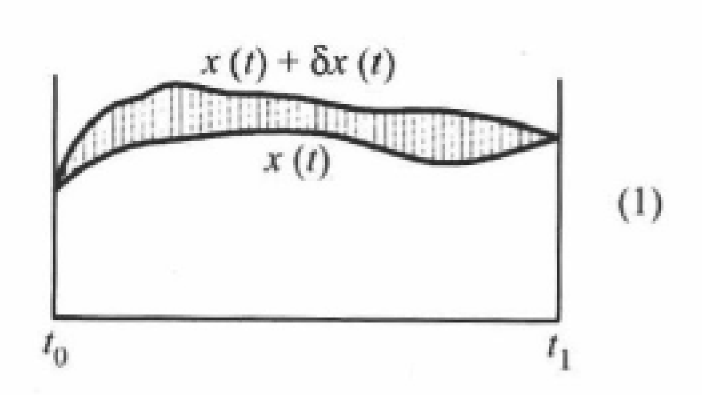
$$[F]_{x(t)}=\frac{\delta f(x)}{\delta x}=\frac{\partial F}{\partial x}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial\dot{x}}.$$

(1) 固定边界条件

通常求泛函的极值时,多数情况为**固定边界**条件,即有

$$\delta x(t_1)=\delta x(t_0)=0,$$

如右图之(1).于是,固定边界下 $f[x(t)]$为平稳的必要条件是其一阶变分导数$[F]_{x(t)}$为零,即



$$[F]_{x(t)}=\frac{\delta f}{\delta x}=\frac{\partial F}{\partial x}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial\dot{x}}=0,$$

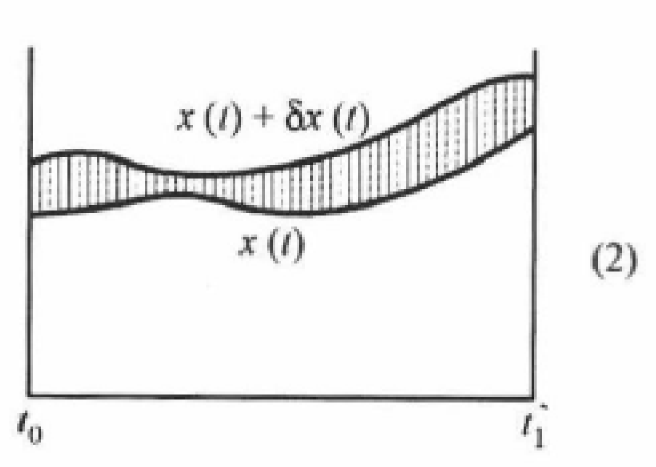
这就是 **Euler[变分]方程.**

(2) 自由边界条件

对于**自由边界**,有

$$\delta x(t_0)\neq 0,\quad \delta x(t_1)\neq 0;$$

如右图之(2).这时 $f[x(t)]$为平稳的必要条件是其一阶变分为零,除要求满足 Euler 方程外,还要求在边界处对变分的贡献为零,即



$$\begin{cases}\dfrac{\partial F}{\partial x}-\dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\partial F}{\partial\dot{x}}=0,\\[2mm] \dfrac{\partial F}{\partial\dot{x}}\bigg|_{t=t_0}=0,\quad \dfrac{\partial F}{\partial\dot{x}}\bigg|_{t=t_1}=0.\end{cases}$$

(3) 变端点问题

当端点 t_1 发生变化时，

$$\begin{aligned}\Delta f &= (f+\Delta f)-f \\ &= \int_{t_0}^{t_1+\Delta t_1} F(\cdots)\mathrm{d}t - \int_{t_0}^{t_1} F(\cdots)\mathrm{d}t \\ &= \int_{t_1}^{t_1+\Delta t_1} F(\cdots)\mathrm{d}t \\ &= F(\cdots)_{t=t_1}\Delta t_1,\end{aligned}$$

其中最后一步用了中值定理. 当端点 t_0 也发生变化时，则有

$$\Delta f = F(\cdots)_{t=t_1}\Delta t_1 - F(\cdots)_{t=t_0}\Delta t_0.$$

所以，在此情况下，泛函的一阶变分应为

$$\delta f[x] = \int_{t_0}^{t_1}\delta x(t)[F]_{x(t)} + \delta x(t_1)\frac{\partial F}{\partial \dot{x}}\bigg|_{t=t_1} - \delta x(t_0)\frac{\partial F}{\partial \dot{x}}\bigg|_{t=t_0} + \delta t_1[F]_{t=t_1} - \delta t_0[F]_{t=t_0}.$$

这时 $f[x(t)]$ 为平稳的必要条件 $\delta f[x]=0$ 除要求满足 Euler 方程及 $\frac{\partial F}{\partial \dot{x}}\Big|_{t=t_1} = \frac{\partial F}{\partial \dot{x}}\Big|_{t=t_0} = 0$ 外，还要求 $[F]_{t=t_1} = [F]_{t=t_0} = 0$.

下面以 Newton 运动方程的推导为例. 取

$$\begin{aligned}F(x,\dot{x},t) &= L(x,\dot{x},t) \\ &= T(\dot{x}) - V(x,t) \\ &= \frac{1}{2}m\dot{x}^2 - V(x,t),\end{aligned}$$

其中 $T(\dot{x}) = \frac{1}{2}m\dot{x}^2$ 为动能，$V(x,t)$ 为势能，而 $L(x,\dot{x},t)$ 则为 **Lagrange 函数**. 由 $\delta\int L\mathrm{d}t = 0$ 在固定边界下的 Euler 变分方程：

$$\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} = 0,$$

得到

$$-\frac{\partial V}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}(m\dot{x}) = 0,$$

即

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\dot{x}) = -\frac{\partial V}{\partial x}$$

或

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2}=-\frac{\partial V}{\partial x},$$

这就是 **Newton 运动方程.**

9.3.3 含高阶导数的情形

对于含高阶导数的泛函,例如

$$f[x(t)]=\int_{t_0}^{t_1}F(x,\dot{x},\ddot{x},\cdots,t)\,\mathrm{d}t,$$

同样可求得其一阶变分为

$$\begin{aligned}\delta f=&\int_{t_0}^{t_1}\delta x(t)[F]_{x(t)}\,\mathrm{d}t\\&+\delta t_1[F]_{t=t_1}-\delta t_0[F]_{t=t_0}\\&+\delta x(t_1)[F]_{\dot{x}(t_1)}-\delta x(t_0)[F]_{\dot{x}(t_0)}\\&+\left(\frac{\mathrm{d}\,\delta x}{\mathrm{d}t}\right)_{t_1}[F]_{\dot{x}(t_1)}-\left(\frac{\mathrm{d}\delta x}{\mathrm{d}t}\right)_{t_0}[F]_{\dot{x}(t_0)}+\cdots;\end{aligned}$$

其中

$$\begin{aligned}[F]_t&=F(x(t),\dot{x}(t),\ddot{x}(t),\cdots,t),\\[F]_{x(t)}&=\frac{\partial F}{\partial x}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dot{x}}+\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial F}{\partial \ddot{x}}-\cdots,\\[F]_{\dot{x}(t)}&=\frac{\partial F}{\partial \dot{x}}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \ddot{x}}+\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial F}{\partial \dddot{x}}-\cdots,\end{aligned}$$

等等. 证明如下：首先,

$$\begin{aligned}x&\to x+\delta x,\\\dot{x}&\to\dot{x}+\frac{\mathrm{d}}{\mathrm{d}t}\delta x,\\\ddot{x}&\to\ddot{x}+\frac{\mathrm{d}^2}{\mathrm{d}t^2}\delta x,\end{aligned}$$

等等;因而

$$\begin{aligned}\delta F=&F\left(x+\delta x,\dot{x}+\frac{\mathrm{d}}{\mathrm{d}t}\delta x,\ddot{x}+\frac{\mathrm{d}^2}{\mathrm{d}t^2}\delta x,\cdots,t\right)-F(x,\dot{x},\ddot{x},\cdots,t)\\=&\frac{\partial F}{\partial x}\delta x+\frac{\partial F}{\partial \dot{x}}\frac{\mathrm{d}}{\mathrm{d}t}\delta x+\frac{\partial F}{\partial \ddot{x}}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\delta x+\cdots.\end{aligned}$$

所以,泛函的一阶变分为

$$\delta f = \delta\int_{t_0}^{t_1} F\mathrm{d}t = \int_{t_0}^{t_1}\delta F\mathrm{d}t + \delta t_1[F]_{t=t_1} - \delta t_0[F]_{t=t_0}.$$

由于

$$\int_{t_0}^{t_1}\delta F\mathrm{d}t = \int_{t_0}^{t_1}\left(\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial \dot{x}}\frac{\mathrm{d}}{\mathrm{d}t}\delta x + \frac{\partial F}{\partial \ddot{x}}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\delta x + \cdots\right)\mathrm{d}t$$

中各项通过分部积分法可得

$$\int_{t_0}^{t_1}\left(\frac{\partial F}{\partial \dot{x}}\frac{\mathrm{d}}{\mathrm{d}t}\delta x\right)\mathrm{d}t = \left(\delta x\frac{\partial F}{\partial \dot{x}}\right)\Big|_{t_0}^{t_1} - \int_{t_0}^{t_1}\delta x\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dot{x}}\right)\mathrm{d}t,$$

$$\int_{t_0}^{t_1}\left(\frac{\partial F}{\partial \ddot{x}}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\delta x\right)\mathrm{d}t = \left(\frac{\mathrm{d}\delta x}{\mathrm{d}t}\frac{\partial F}{\partial \ddot{x}}\right)\Big|_{t_0}^{t_1} - \left(\delta x\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \ddot{x}}\right)\Big|_{t_0}^{t_1} + \int_{t_0}^{t_1}\delta x\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial F}{\partial \ddot{x}}\right)\mathrm{d}t,$$

$$\int_{t_0}^{t_1}\left(\frac{\partial F}{\partial \dddot{x}}\frac{\mathrm{d}^3}{\mathrm{d}t^3}\delta x\right)\mathrm{d}t = \left(\frac{\mathrm{d}^2\delta x}{\mathrm{d}t^2}\frac{\partial F}{\partial \dddot{x}}\right)\Big|_{t_0}^{t_1} - \left(\frac{\mathrm{d}\delta x}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dddot{x}}\right)\Big|_{t_0}^{t_1}$$
$$+ \left(\delta x\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial F}{\partial \dddot{x}}\right)\Big|_{t_0}^{t_1} - \int_{t_0}^{t_1}\delta x\left(\frac{\mathrm{d}^3}{\mathrm{d}t^3}\frac{\partial F}{\partial \dddot{x}}\right)\mathrm{d}t,$$

等等. 于是,合并后即可得前述结果.

9.3.4 几个变函数的情形

现在推广到有几个变函数的情形,

$$f[x(t), y(t), \cdots] = \int_a^b F(x(t), y(t), \cdots; \dot{x}(t), \dot{y}(t), \cdots; \cdots; t)\mathrm{d}t,$$

容易求得泛函的一阶变分为

$$\delta f = \int_a^b \{\delta x(t)[F]_{x(t)} + \delta y(t)[F]_{y(t)} + \cdots\}\mathrm{d}t + (\cdots)\Big|_a^b,$$

其中$(\cdots)\Big|_a^b$代表边界效应,而$[F]_{x(t)}$,$[F]_{y(t)}$为变分导数,其表达式是

$$[F]_{x(t)} = \frac{\delta F}{\delta x} = \frac{\partial F}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dot{x}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial F}{\partial \ddot{x}} - \cdots,$$

$$[F]_{y(t)} = \frac{\delta F}{\delta y} = \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dot{y}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial F}{\partial \ddot{y}} - \cdots,$$

等等. 于是,固定边界下,泛函 $f[x(t), y(t), \cdots]$平稳的必要条件是

$$[F]_{x(t)} = 0, \quad [F]_{y(t)} = 0, \quad \cdots.$$

9.3.5 变函数为复函数的情形

函数中出现虚数 i 时称为复函数. 例如,量子力学中的波函数 $\psi(x)$就是 x 的复函数. 应注意,复函数与复变函数不同,因为后者的自变量 z 为复数,而前者的自变

量 x 仍为实数.

因为复函数 $\psi(x)$ 总可通过展开而将其表示为 $\psi(x)=\varphi_1(x)+\mathrm{i}\varphi_2(x)$，其复共轭为 $\psi^*(x)=\varphi_1(x)-\mathrm{i}\varphi_2(x)$. 所以，变函数为复函数的情形实际上相当于两个变函数的情形，$\delta\psi^*$ 和 $\delta\psi$ 可独立变分(因为 $\delta\varphi_1$ 和 $\delta\varphi_2$ 可独立变分).

例如，$\delta\int\mathscr{L}(\psi,\psi^*)\mathrm{d}x=0$ 的 Euler 方程为$\dfrac{\partial\mathscr{L}}{\partial\psi}=0$ 和$\dfrac{\partial\mathscr{L}}{\partial\psi^*}=0$. 由于物理量 $\mathscr{L}$ 总是实值，因而这两个方程是共轭的，取其中之一即可.

9.3.6 几个参变量的情形

下面再来考虑几个参变量的情形，

$$f[x(u,v)]=\iint_R F(x;x_u,x_v;u,v)\mathrm{d}u\mathrm{d}v,$$

其中 $x_u=\dfrac{\partial x}{\partial u},x_v=\dfrac{\partial x}{\partial v}$；$R$ 为积分域.

因为 $\delta(x_u)=\dfrac{\partial}{\partial u}\delta x,\dfrac{\partial}{\partial u}(ab)-a\dfrac{\partial b}{\partial u}=b\dfrac{\partial a}{\partial u}$，所以有

$$\begin{aligned}\delta f[x(u,v)]&=\iint_R\left\{\frac{\partial F}{\partial x}\delta x+\frac{\partial F}{\partial x_u}\left(\frac{\partial}{\partial u}\delta x\right)+\frac{\partial F}{\partial x_v}\left(\frac{\partial}{\partial v}\delta x\right)\right\}\mathrm{d}u\mathrm{d}v\\&=\iint_R\left\{\frac{\partial F}{\partial x}-\frac{\partial}{\partial u}\left(\frac{\partial F}{\partial x_u}\right)-\frac{\partial}{\partial v}\left(\frac{\partial F}{\partial x_v}\right)\right\}\delta x\mathrm{d}u\mathrm{d}v\\&\quad+\iint_R\left\{\frac{\partial}{\partial u}\left(\delta x\frac{\partial F}{\partial x_u}\right)+\frac{\partial}{\partial v}\left(\delta x\frac{\partial F}{\partial x_v}\right)\right\}\mathrm{d}u\mathrm{d}v.\end{aligned}$$

另外，将 Stokes 定理 $\int_R(\nabla\times\boldsymbol{H})\cdot\boldsymbol{n}\mathrm{d}S=\oint_C\boldsymbol{H}\cdot\mathrm{d}\boldsymbol{l}$ 应用于平面，若 $\boldsymbol{H}=\boldsymbol{e}_1H_1+\boldsymbol{e}_2H_2$，$\mathrm{d}S=\mathrm{d}u\mathrm{d}v,\mathrm{d}\boldsymbol{l}=\boldsymbol{e}_1\mathrm{d}u+\boldsymbol{e}_2\mathrm{d}v,\boldsymbol{n}=\boldsymbol{e}_3$，得到①

$$\iint_R\left(\frac{\partial}{\partial u}H_2-\frac{\partial}{\partial v}H_1\right)\mathrm{d}u\mathrm{d}v=\oint_C(H_1\mathrm{d}u+H_2\mathrm{d}v).$$

现在令 $H_1=-\delta x\dfrac{\partial F}{\partial x_v}$，$H_2=\delta x\dfrac{\partial F}{\partial x_u}$，可使 δf 的后一积分变为 $\oint_C\delta x(u,v)\times\left(\dfrac{\partial F}{\partial x_u}\mathrm{d}v-\dfrac{\partial F}{\partial x_v}\mathrm{d}u\right)$，$C$ 为积分域 R 的周线. 它代表边界贡献部分. 于是，最后可求得泛函 $f[x(u,v)]$的一阶变分为

---

① 此结果一般称为 Green 公式，通常表述为(S 为闭合曲线所围之面积)

$$\oint_C(P\mathrm{d}x+Q\mathrm{d}y)=\iint_S\left[\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right]\mathrm{d}x\mathrm{d}y.$$

$$\delta f[x(u,v)]=\iint_R \delta x(u,v)[F]_{x(u,v)}\mathrm{d}u\mathrm{d}v+\oint_C \delta x(u,v)\left(\frac{\partial F}{\partial x_u}\mathrm{d}v-\frac{\partial F}{\partial x_v}\mathrm{d}u\right),$$

其中泛函导数$[F]_{x(u,v)}$为

$$[F]_{x(u,v)}=\frac{\delta f}{\delta x}=\frac{\partial F}{\partial x}-\frac{\partial}{\partial u}\frac{\partial F}{\partial x_u}-\frac{\partial}{\partial v}\frac{\partial F}{\partial x_v}.$$

对于自由边界而言，泛函 $f[x(u,v)]$是平稳的必要条件是

$$\begin{cases}[F]_{x(u,v)}=0,\\ \dfrac{\partial F}{\partial x_u}\mathrm{d}v-\dfrac{\partial F}{\partial x_v}\mathrm{d}u=0;\end{cases}$$

而对固定边界而言，只需$[F]_{x(u,v)}=0$.

另一方面，若泛函中含有高阶导数，则其变分导数应修正为

$$[F]_{x(u,v)}=\frac{\partial F}{\partial x}-\frac{\partial}{\partial u}\frac{\partial F}{\partial x_u}-\frac{\partial}{\partial v}\frac{\partial F}{\partial x_v}+\frac{\partial^2}{\partial u^2}\frac{\partial F}{\partial x_{uu}}+2\frac{\partial^2}{\partial u\partial v}\frac{\partial F}{\partial x_{uv}}+\frac{\partial^2}{\partial v^2}\frac{\partial F}{\partial x_{vv}}-\cdots;$$

$[F]_{x(u,v)}=0$ 仍是泛函 $f[x(u,v)]$在固定边界下的平稳条件. 对于自由边界而言，此外还需边界对变分贡献为零等若干补充条件，这里从略.

9.4 Ritz 方法

对于泛函的极值问题，有时可以采用**直接法**去求解. 因为任意函数 $x(t)$均可用函数空间中的一组基函数 $\varphi_1(t),\cdots,\varphi_n(t),\cdots$予以展开：

$$x(t)=\sum_{n=1}^{\infty}c_n\varphi_n(t);$$

于是，泛函 $f[x(t)]$用上述展开式代入后得

$$f[x(t)]=\int_a^b F(x,\dot{x},\cdots,t)\mathrm{d}t=\int_a^b F\left(\left\{\sum c_n\varphi_n\right\},\left\{\sum c_n\dot{\varphi}_n\right\},\cdots,t\right)\mathrm{d}t,$$

完成积分后可将之表达为$\{c_n\}$的函数，即

$$f[x(t)]=\mathscr{F}(\{c_n\}).$$

因而泛函 $f[x(t)]$的极值问题化为函数 $\mathscr{F}(\{c_n\})$的极值问题，即，$\delta f=0$ 导致要求

$$\frac{\partial\mathscr{F}}{\partial c_k}=0,\quad k=1,2,\cdots,n,\cdots.$$

Ritz 方法就是要选择（“猜”）一个适当的“坐标系”$\varphi_n(t)$，要选得好（要“会猜”），使 $\sum c_n\varphi_n(t)$ 能很快地收敛，以致只要很少的几项就能很好地代表 $x(t)$，用它去算出 $f[x(t)]$，即 $\mathscr{F}(c_1,c_2,\cdots)$，再由上述代数方程组$\frac{\partial\mathscr{F}}{\partial c_k}=0(k=1,2,\cdots)$来确定使泛函为极值的此几项系数 c_k 之值. 应注意，应用 Ritz 方法时，实际上算起来只是近似，但由于无需求解微分方程(Euler 方程)，实用上很方便；所以，它是很有用的方法.

9.5 条件极值问题

9.5.1 函数的条件极值

众所周知,要在附加条件 $g_i(x_1,\cdots,x_n)=0(i=1,\cdots,m)$ 下求函数 $f(x_1,\cdots,x_n)$的极值,只需引入 **Lagrange 乘子** $\lambda_i(i=1,\cdots,m)$,再来求 $f^*=f+\sum_i\lambda_i g_i$ 作为 $\{x_k\}$, $\{\lambda_i\}$函数的无条件极值即可. 也就是说,函数 f 的条件极值点 x_k 及相应 Lagrange 乘子 λ_i 可由

$$\frac{\partial f^*}{\partial x_k}=\frac{\partial f}{\partial x_k}+\sum_i\lambda_i\frac{\partial g_i}{\partial x_k}=0 \quad (k=1,\cdots,n),$$

$$\frac{\partial f^*}{\partial \lambda_i}=g_i=0 \quad (i=1,\cdots,m)$$

予以确定.

9.5.2 泛函的条件极值

对于泛函,也可类似地引入 **Lagrange 乘子**来求其**条件极值问题**.

例如,在归一化条件 $\int\{\varphi(x)\}^2\mathrm{d}x=1$ 下求泛函

$$f[\varphi(x)]=\iint\varphi(x)K(x,y)\varphi(y)\mathrm{d}x\mathrm{d}y\equiv\iint F\mathrm{d}x\mathrm{d}y$$

的极值;可令 $g[\varphi(x)]\equiv\int\{\varphi(x)\}^2\mathrm{d}x-1=0$,引入 Lagrange 乘子 λ,而变为求新泛函

$$f^*[\varphi(x)]=f[\varphi(x)]-\lambda g[\varphi(x)]$$

的无条件极值. 将新泛函 f^* 看做含参变量 λ 的泛函,其一阶变分为

$$\delta f^*[\varphi(x)](\lambda)=\int\delta\varphi(x)[F]_{\varphi(x)}\mathrm{d}x-\lambda\int\delta\varphi(x)[G]_{\varphi(x)}\mathrm{d}x-(\delta\lambda)g[\varphi];$$

这里 $G=\{\varphi(x)\}^2-\delta(x)$. 于是,$\delta f^*=0$ 给出其平稳条件为

$$\begin{cases}[F^*]_{\varphi(x)}=[F]_{\varphi(x)}-\lambda[G]_{\varphi(x)}=0,\\ \dfrac{\delta f^*}{\delta\lambda}=-g[\varphi]=0.\end{cases}$$

不失一般性,假定 $K(x,y)$是对称的,可得

$$[F]_{\varphi(x)}=2\int K(x,y)\varphi(y)\mathrm{d}y,\quad [G]_{\varphi(x)}=2\varphi(x);$$

上述平稳条件可明确写出为

$$\begin{cases} \int K(x,y)\varphi(y)\mathrm{d}y = \lambda\varphi(x), \\ \int \{\varphi(x)\}^2\mathrm{d}x = 1. \end{cases}$$

这个结果表明：后者即原归一化条件；而前者变成本征值问题，Lagrange 乘子 λ 即其本征值. 还可注意到，此本征值 λ 即为原泛函 $f[\varphi(x)]$ 的极值，因为对前式进行 $\int \varphi(x)[\cdots]\mathrm{d}x$ 运算立即得到

$$\iint \varphi(x)K(x,y)\varphi(y)\mathrm{d}x\mathrm{d}y = \lambda\int\{\varphi(x)\}^2\mathrm{d}x = \lambda.$$

最后还应指出，这个条件极值问题与下列泛函

$$\frac{\iint \varphi(x)K(x,y)\varphi(y)\mathrm{d}x\mathrm{d}y}{\int\{\varphi(x)\}^2\mathrm{d}x}$$

的极值问题相当.

9.5.3 一般的条件极值

现在讨论一般的条件极值问题.

例如，在附加条件 $g_1=0, g_2=0, g(t)=0,\cdots$ 下求 f 的极值. 这里 f 可以是函数 $f(x_1,x_2)$、复合函数 $f(x(t))$、泛函 $f[x(t)]$ 或带参变量的泛函 $f[x(t)](x_1,x_2)$；而附加条件则除了可以是上述各种形式外，还可包含变量的微分或积分形式，具体如 $g(t)=G(x(t),\dot{x}(t),\cdots)=0$，以及诸如 $x(0)-a=0$，$\dot{x}(1)-\beta=0$，等等.

求极值的步骤依然是引入 Lagrange 乘子，令

$$f^* = f+\lambda_1 g_1+\lambda_2 g_2+\int_a^b \lambda(t)g(t)\mathrm{d}t,$$

在求一阶变分 $\delta f^*=0$ 时，将 $\lambda_1,\lambda_2,\lambda(t)$ 与 $x_1,x_2,x(t)$ 等一起作为独立变量来对待. 应该指出，这里的 λ_1,λ_2 是数而 $\lambda(t)$ 则是 t 的函数，所以有时候称前者为 **Lagrange 乘数**，称后者为 **Lagrange 乘函**.

9.6 曲线坐标系下 Laplace 方程的推导

现在利用变分法来推导在曲线坐标系下的 Laplace 方程：$\nabla^2 u=0$.

9.6.1 变分法问题

首先将此化为变分法问题. 考虑下列泛函：

$$f[u(x,y,z)]=\iiint_V\left\{\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2\right\}\mathrm{d}x\mathrm{d}y\mathrm{d}z=\int_V F(u_x,u_y,u_z)\mathrm{d}V,$$

固定边界下的一阶变分是

$$\begin{aligned}\delta f&=\int_V \delta F\mathrm{d}V\\&=\int_V\left\{2\left(\frac{\partial u}{\partial x}\right)\frac{\partial}{\partial x}\delta u+2\left(\frac{\partial u}{\partial y}\right)\frac{\partial}{\partial y}\delta u+2\left(\frac{\partial u}{\partial z}\right)\frac{\partial}{\partial z}\delta u\right\}\mathrm{d}V\\&=2\int_V(\nabla u)\cdot(\nabla\delta u)\mathrm{d}V,\end{aligned}$$

这里应用了 $\delta\nabla u=\nabla\delta u$.

由矢量导数运算公式

$$\nabla\cdot(\varphi\boldsymbol{v})=\boldsymbol{v}\cdot\nabla\varphi+\varphi\nabla\cdot\boldsymbol{v},$$

令 $\boldsymbol{v}=\nabla u,\varphi=\delta u$,并注意到$\nabla\cdot(\nabla u)=\nabla^2 u$,则 δf 可化为

$$\delta f=2\int_V(\nabla u)\cdot(\nabla\delta u)\mathrm{d}V=2\int_V\nabla\cdot(\delta u\,\nabla u)\mathrm{d}V-2\int_V\delta u\,\nabla^2 u\mathrm{d}V.$$

下面应用 Gauss 散度定理：

$$\int_V(\nabla\cdot\boldsymbol{A})\mathrm{d}V=\oint_S\boldsymbol{A}\cdot\boldsymbol{n}\mathrm{d}S,$$

将体积分化为面积分,这里 S 是包围体积 V 的闭曲面,$\boldsymbol{n}$ 为曲面 S 上的外法线方向的单位矢量. 为此令 $\boldsymbol{A}=\delta u\,\nabla u$,则 δf 可化为

$$\delta f=2\oint_S\delta u(\nabla u)\cdot\boldsymbol{n}\mathrm{d}S-2\int_V\delta u\,\nabla^2 u\mathrm{d}V.$$

若假定在曲面 S 上有 $\delta u=0$,一阶变分 δf 最后化为

$$\delta f=-2\int_V\delta u\cdot\nabla^2 u\mathrm{d}V,$$

因而,当且仅当 $\nabla^2 u=0$ 时有 $\delta f=0$. 换句话说,

$$\nabla^2 u=0\iff\delta\int_V(u_x^2+u_y^2+u_z^2)\mathrm{d}V=0,$$

即,二者是等价关系. 所以,想要寻求 Laplace 方程的普遍公式,可由上述变分法问题出发来推导.

9.6.2 坐标变换

现在来进行坐标变换,要使直角坐标系下的 Laplace 方程

$$\nabla^2 u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=0,$$

变换为曲线坐标系下的普遍形式.

考虑下列一般坐标变换

$$\left.\begin{aligned} x &= x(\xi_1,\xi_2,\xi_3) \\ y &= y(\xi_1,\xi_2,\xi_3) \\ z &= z(\xi_1,\xi_2,\xi_3) \end{aligned}\right\} \xrightarrow{\text{求解得}} \left\{\begin{aligned} \xi_1 &= \xi_1(x,y,z) \\ \xi_2 &= \xi_2(x,y,z) \\ \xi_3 &= \xi_3(x,y,z) \end{aligned}\right.$$

直角坐标系		曲线坐标系
(x,y,z)		(ξ_1,ξ_2,ξ_3)
$u(x,y,z)$		$u(\xi_1,\xi_2,\xi_3)$
$u_x^2+u_y^2+u_z^2$	$\longrightarrow$	$\sum\limits_{i,k} g^{ik}u_iu_k$
$u_x=\dfrac{\partial u}{\partial x},\cdots$		$u_i=\dfrac{\partial u}{\partial \xi_i}$

$$g^{ik}=\frac{\partial \xi_i}{\partial x}\frac{\partial \xi_k}{\partial x}+\frac{\partial \xi_i}{\partial y}\frac{\partial \xi_k}{\partial y}+\frac{\partial \xi_i}{\partial z}\frac{\partial \xi_k}{\partial z}$$

$$\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2 \longrightarrow \sum_{i,k} g_{ik}\,\mathrm{d}\xi_i\,\mathrm{d}\xi_k$$

$$g_{ik}=\frac{\partial x}{\partial \xi_i}\frac{\partial x}{\partial \xi_k}+\frac{\partial y}{\partial \xi_i}\frac{\partial y}{\partial \xi_k}+\frac{\partial z}{\partial \xi_i}\frac{\partial z}{\partial \xi_k}$$

$$\sum_i g_{ik}g^{il}=\delta_{kl}$$

在直角坐标系中，

$$\nabla^2 u=0 \iff \delta\int_V F\,\mathrm{d}V=0,$$

$$F=u_x^2+u_y^2+u_z^2.$$

现在进行坐标变换$(x,y,z)\to(\xi_1,\xi_2,\xi_3)$，

$$\delta\int F\mathrm{d}V\to\delta\int\Big\{\sum_{i,k}g^{ik}u_iu_k\Big\}\frac{\partial(x,y,z)}{\partial(\xi_1,\xi_2,\xi_3)}\mathrm{d}\xi_1\,\mathrm{d}\xi_2\,\mathrm{d}\xi_3,$$

其中

$$\frac{\partial(x,y,z)}{\partial(\xi_1,\xi_2,\xi_3)}=\begin{vmatrix} \dfrac{\partial x}{\partial \xi_1} & \dfrac{\partial x}{\partial \xi_2} & \dfrac{\partial x}{\partial \xi_3} \\ \dfrac{\partial y}{\partial \xi_1} & \dfrac{\partial y}{\partial \xi_2} & \dfrac{\partial y}{\partial \xi_3} \\ \dfrac{\partial z}{\partial \xi_1} & \dfrac{\partial z}{\partial \xi_2} & \dfrac{\partial z}{\partial \xi_3} \end{vmatrix}.$$

于是，在曲线坐标系中，变成了

$$\nabla^2 u=0 \iff \delta\int G\mathrm{d}\tau=0,$$

其中

$$G=\Big\{\sum_{i,k}g^{ik}u_iu_k\Big\}\frac{\partial(x,y,z)}{\partial(\xi_1,\xi_2,\xi_3)},\quad \mathrm{d}\tau=\mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}\xi_3.$$

9.6.3 一般结果

现在来求曲线坐标系下 Laplace 方程的一般形式.

注意到,根据行列式的性质有

$$\left\{\frac{\partial(x,y,z)}{\partial(\xi_1,\xi_2,\xi_3)}\right\}^2=\left\{\frac{\partial(x,y,z)}{\partial(\xi_1,\xi_2,\xi_3)}\right\}^{\mathrm{T}}\left\{\frac{\partial(x,y,z)}{\partial(\xi_1,\xi_2,\xi_3)}\right\}$$

$$=\det|g_{ik}|=\frac{1}{\det|g^{ik}|}=g,$$

$$\frac{\partial(x,y,z)}{\partial(\xi_1,\xi_2,\xi_3)}=\sqrt{g}.$$

于是,

$$f[u(\xi_1,\xi_2,\xi_3)]=\int\Big\{\sum_{i,k}g^{ik}\sqrt{g}u_iu_k\Big\}\mathrm{d}\tau=\int G\mathrm{d}\tau;$$

$$\delta\int G\mathrm{d}\tau=\delta\int\Big\{\sum_{i,k}g^{ik}\sqrt{g}u_iu_k\Big\}\mathrm{d}\tau=0\rightarrow$$

$$[G]_{u(\xi_1,\xi_2,\xi_3)}=\frac{\partial G}{\partial u}-\sum_i\frac{\partial}{\partial\xi_i}\frac{\partial G}{\partial u_i}=-\sum_i\frac{\partial}{\partial\xi_i}\frac{\partial G}{\partial u_i}$$

$$=-2\sum_{i,k}\frac{\partial}{\partial\xi_i}\Big(g^{ik}\sqrt{g}\frac{\partial u}{\partial\xi_k}\Big)=0.$$

因为

$$\delta f=-2\int\delta u\,\nabla^2u\mathrm{d}V=-2\int\delta u\,\nabla^2u\sqrt{g}\,\mathrm{d}\tau\rightarrow$$

$$\delta\int G\mathrm{d}\tau=\int\delta u[G]_{u(\xi_1,\xi_2,\xi_3)}\mathrm{d}\tau=-2\int\delta u\Big\{\sum_{i,k}\frac{\partial}{\partial\xi_i}\Big(g^{ik}\sqrt{g}\frac{\partial u}{\partial\xi_k}\Big)\Big\}\mathrm{d}\tau;$$

所以,最后得到 **Laplace 方程的一般形式为**

$$\boxed{\nabla^2u=\frac{1}{\sqrt{g}}\sum_{i,k}\frac{\partial}{\partial\xi_i}\Big(g^{ik}\sqrt{g}\frac{\partial u}{\partial\xi_k}\Big)=0\,.}$$

应该注意到,它与张量运算中所得广义结果的一致性[见 5.11.4 节之(3)].

9.6.4 具体例子

(1) 球面坐标系

$$(x,y,z)\rightarrow(\xi_1,\xi_2,\xi_3)=(r,\theta,\varphi),$$

$$\begin{cases}x=r\sin\theta\cos\varphi,\\ y=r\sin\theta\sin\varphi,\\ z=r\cos\theta;\end{cases}$$

$$dx^2+dy^2+dz^2=dr^2+r^2d\theta^2+r^2\sin^2\theta d\varphi^2,$$

$$g_{11}=1,\quad g_{22}=r^2,\quad g_{33}=r^2\sin^2\theta;$$

$$g^{11}=1,\quad g^{22}=\frac{1}{r^2},\quad g^{33}=\frac{1}{r^2\sin^2\theta};$$

$$\sqrt{g}=\sqrt{1\cdot r^2\cdot r^2\sin^2\theta}=r^2\sin\theta,$$

$$dv=\sqrt{g}d\tau,\quad 即\quad dxdydz=r^2\sin\theta drd\theta d\varphi.$$

$$\nabla^2u=\frac{1}{r^2\sin\theta}\left\{\frac{\partial}{\partial r}\left(r^2\sin\theta\frac{\partial u}{\partial r}\right)+\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right)+\frac{\partial}{\partial\varphi}\left(\frac{1}{\sin\theta}\frac{\partial u}{\partial\varphi}\right)\right\},$$

化简后最后得到

$$\nabla^2u=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2u}{\partial\varphi^2}.$$

(2) 柱面坐标系

$$(x,y,z)\to(\rho,\varphi,z),$$

$$\begin{cases}x=\rho\cos\varphi,\\ y=\rho\sin\varphi,\\ z=z;\end{cases}$$

$$dx^2+dy^2+dz^2=d\rho^2+\rho^2d\varphi^2+dz^2,$$

$$g_{11}=1,\quad g_{22}=\rho^2,\quad g_{33}=1;$$

$$g^{11}=1,\quad g^{22}=\frac{1}{\rho^2},\quad g^{33}=1.$$

$$\sqrt{g}=\rho,\quad dxdydz=\rho d\rho d\varphi dz,$$

$$\nabla^2u=\frac{1}{\rho}\left\{\frac{\partial}{\partial\rho}\left(\rho\frac{\partial u}{\partial\rho}\right)+\frac{\partial}{\partial\varphi}\left(\frac{1}{\rho}\frac{\partial u}{\partial\varphi}\right)+\frac{\partial}{\partial z}\left(\rho\frac{\partial u}{\partial z}\right)\right\},$$

$$\nabla^2u=\frac{\partial^2u}{\partial\rho^2}+\frac{1}{\rho}\frac{\partial u}{\partial\rho}+\frac{1}{\rho^2}\frac{\partial^2u}{\partial\varphi^2}+\frac{\partial^2u}{\partial z^2}.$$

由于所讨论问题常具有特定对称性，因而采用特定坐标系来处理会比较方便. 例如，讨论无穷长荷电导线的电势，采用柱面坐标较方便；而求宽度为 $2a$ 的无穷长荷电薄片的电势，或者，求流经平面障板中宽度为 $2a$ 的狭缝之不可压缩无黏性流体的速度势，则以采用椭圆柱坐标较方便. 还有抛物柱坐标也是较有用的，下面分别作简要介绍.

(3) 椭圆柱坐标系

$$\begin{cases}x=a\,\mathrm{ch}\,\xi_1\cos\xi_2,\\ y=a\,\mathrm{sh}\,\xi_1\sin\xi_2,\\ z=z.\end{cases}$$

因此，在 $z=\mathrm{const}$ 的平面中，

$$\left(\frac{x}{a\ \mathrm{ch}\,\xi_1}\right)^2+\left(\frac{y}{a\ \mathrm{sh}\,\xi_1}\right)^2=1\quad(\text{共焦椭圆},\xi_1=\mathrm{const});$$

$$\left(\frac{x}{a\cos\xi_2}\right)^2-\left(\frac{y}{a\sin\xi_2}\right)^2=1\quad(\text{共焦双曲线},\xi_2=\mathrm{const}).$$

$$\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2=a^2(\mathrm{ch}^2\xi_1-\cos^2\xi_2)(\mathrm{d}\xi_1^2+\mathrm{d}\xi_2^2)+\mathrm{d}z^2,$$

$$\nabla^2u=\frac{1}{a^2(\mathrm{ch}^2\xi_1-\cos^2\xi_2)}\left\{\frac{\partial^2u}{\partial\xi_1^2}+\frac{\partial^2u}{\partial\xi_2^2}\right\}+\frac{\partial^2u}{\partial z^2}.$$

另一种椭圆柱坐标是 ξ,η,z:

$$\xi=\mathrm{ch}\,\xi_1,\quad \eta=\cos\xi_2;$$

$$x=a\xi\eta,\quad y=a\sqrt{(\xi^2-1)(1-\eta^2)},\quad z=z.$$

$$\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2=a^2\left\{\frac{\xi^2-\eta^2}{\xi^2-1}\mathrm{d}\xi^2+\frac{\xi^2-\eta^2}{1-\eta^2}\mathrm{d}\eta^2\right\}+\mathrm{d}z^2;$$

$$\nabla^2u=\frac{1}{a^2(\xi^2-\eta^2)}\left\{\sqrt{\xi^2-1}\frac{\partial}{\partial\xi}\left(\sqrt{\xi^2-1}\frac{\partial u}{\partial\xi}\right)+\sqrt{1-\eta^2}\frac{\partial}{\partial\eta}\left(\sqrt{1-\eta^2}\frac{\partial u}{\partial\eta}\right)\right\}+\frac{\partial^2u}{\partial z^2}.$$

(4) 抛物柱坐标系

$$\begin{cases}x=\dfrac{1}{2}(\xi_2^2-\xi_1^2),\\ y=\xi_1\xi_2,\\ z=z.\end{cases}$$

因此,在 z=const 的平面中,有两组互相正交的抛物线,其焦点位于 z 轴上,

$$y^2=\xi_1^2(\xi_1^2+2x)\quad(\xi_1=\mathrm{const}),$$

$$y^2=\xi_2^2(\xi_2^2-2x)\quad(\xi_2=\mathrm{const}).$$

$$\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2=(\xi_1^2+\xi_2^2)(\mathrm{d}\xi_1^2+\mathrm{d}\xi_2^2)+\mathrm{d}z^2,$$

$$\nabla^2u=\frac{1}{\xi_1^2+\xi_2^2}\left\{\frac{\partial^2u}{\partial\xi_1^2}+\frac{\partial^2u}{\partial\xi_2^2}\right\}+\frac{\partial^2u}{\partial z^2}.$$

9.7 变分原理

9.7.1 Hamilton 正则运动方程

前面曾经应用变分法推导出质点一维运动的 Newton 方程.不难看出,对于质点系,若引入 **Lagrange 函数**

$$L=T-V=L(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_n;\dot{\boldsymbol{r}}_1,\cdots,\dot{\boldsymbol{r}}_n;t),$$

则由 $\delta\int_{t_0}^{t_1} L\mathrm{d}t = 0$ 立即得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\boldsymbol{r}}_i}\right) - \frac{\partial L}{\partial \boldsymbol{r}_i} = 0 \quad (i = 1, \cdots, n);$$

此即质点系的 **Newton 运动方程.**

然而,作为力学系,若质点系的结构使其位形可仅由 g 个**广义坐标** $q_a(a=1, \cdots, g)$ 来描述,则其 Lagrange 函数可表达为 $L(q_1, \cdots, q_g; \dot{q}_1, \cdots, \dot{q}_g; t)$, 则由变分法 $\delta\int_{t_0}^{t_1} L\mathrm{d}t = 0$ 给出 Euler 方程:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_a}\right) - \frac{\partial L}{\partial q_a} = 0 \quad (a = 1, \cdots, g);$$

这是 g 个二阶常微分方程,称为力学系的 **Lagrange 运动方程.** 因此,Euler 变分方程又常称为 **Euler-Lagrange 方程.**

现在引入**广义动量**

$$p_a = \frac{\partial L}{\partial \dot{q}_a} \quad (a = 1, \cdots, g),$$

并由此反解出 $\dot{q}_a$,则

$$\dot{q}_a = \dot{q}_a(q_1, \cdots, q_g; p_1, \cdots, p_g; t);$$

同时,进行 **Legendre 变换:**

$$H = \sum_b p_b \dot{q}_b - L = H(q_1, \cdots, q_g; p_1, \cdots, p_g; t),$$

这样引入的函数 H 称为系统的 **Hamilton 函数.** 而广义坐标 $q_1, \cdots, q_g$ 与广义动量 $p_1, \cdots, p_g$ 一起则称为**正则变量.**

于是,

$$\frac{\partial H}{\partial q_a} = \sum_b \left(p_b - \frac{\partial L}{\partial \dot{q}_b}\right)\frac{\partial \dot{q}_b}{\partial q_a} - \frac{\partial L}{\partial q_a} = -\frac{\partial L}{\partial q_a},$$

$$\frac{\partial H}{\partial p_a} = \dot{q}_a + \sum_b \left(p_b - \frac{\partial L}{\partial \dot{q}_b}\right)\frac{\partial \dot{q}_b}{\partial p_a} = \dot{q}_a,$$

求和项消去是由于广义动量的定义 $p_b = \dfrac{\partial L}{\partial \dot{q}_b}$;另外,由于 Lagrange 运动方程

$$\frac{\partial L}{\partial q_a} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_a} = \frac{\mathrm{d}}{\mathrm{d}t}p_a = \dot{p}_a,$$

最后得到 $2g$ 个一阶常微分方程

$$\dot{q}_a = \frac{\partial H}{\partial p_a}, \quad \dot{p}_a = -\frac{\partial H}{\partial q_a} \quad (a = 1, \cdots, g);$$

它们称为 **Hamilton 正则运动方程**,简称**正则方程.** 应该指出,它们也可由变分法

$$\delta\int_{t_0}^{t_1}\left\{\sum_b p_b\dot{q}_b - H\right\}\mathrm{d}t = 0$$

直接导出.换句话说,正则方程与上述变分法问题是等价的.

另外,选择任一函数 G,通过

$$\sum_a p_a \mathrm{d}q_a - H\mathrm{d}t = \sum_a P_a \mathrm{d}Q_a - K\mathrm{d}t + \mathrm{d}G$$

所进行的变换称为**正则变换**;因为,由变分法

$$0=\delta\int_{t_0}^{t_1}\left\{\sum p_a\dot{q}_a - H\right\}\mathrm{d}t = \delta\int_{t_0}^{t_1}\left\{\sum P_a\dot{Q}_a - K\right\}\mathrm{d}t + \delta G\Big|_{t_0}^{t_1}$$

可知,

$$\dot{Q}_a = \frac{\partial K}{\partial P_a},\quad \dot{P}_a = -\frac{\partial K}{\partial Q_a}\quad (a = 1,\cdots,g);$$

也就是说,经过正则变换后,$\{Q_a,P_a\}$为新正则变量,K 为新 Hamilton 函数,它们使新的正则方程保持形式不变.

特别是,若所选之 G 使 $K\equiv 0$,即,如果 G 是 **Hamilton-Jacobi 偏微分方程**

$$H\left(q_a,\frac{\partial G}{\partial q_a},t\right)+\frac{\partial G}{\partial t} = 0$$

的某个完全解 $G=G(q_1,\cdots,q_g;\alpha_1,\cdots,\alpha_g;t)$,则显然有

$$Q_a = \alpha_a,\quad P_a = \beta_a\quad (a = 1,\cdots,g)$$

均为常量.同时,亦可反解求得原正则方程之解 q_a 和 p_a 为此二组常量$\{\alpha_a\}$,$\{\beta_a\}$及 t 的函数.换言之,Hamilton 正则方程与 Hamilton-Jacobi 偏微分方程等价.

尽管一般认为偏微分方程比常微分方程组更难解,然而在处理力学问题时,经常是通过求解 Hamilton-Jacobi 方程来求得正则方程之解.这是由于 Hamilton-Jacobi 方程并不依赖于未知函数本身而仅依赖于其一阶偏导数,此特征对方程的求解有很大帮助.

9.7.2 Maxwell 电磁场方程组

对于电场强度为 $\boldsymbol{E}$、磁感应强度为 $\boldsymbol{B}$ 的电磁场,可知(用≙表示相当于)

$$\int \frac{1}{2}\varepsilon E^2\mathrm{d}\boldsymbol{x} = \text{总电能} \triangleq \text{势能},$$

$$\int \frac{1}{2}\frac{B^2}{\mu}\mathrm{d}\boldsymbol{x} = \text{总磁能} \triangleq \text{动能},$$

其中 ε 为电容率,μ 为磁导率.若引进 **Lagrange 密度函数** $L= \frac{1}{2}\left(\frac{1}{\mu}B^2-\varepsilon E^2\right)$,可得磁场强度 $\boldsymbol{H}$ 和电位移 $\boldsymbol{D}$ 分别为

$$\boldsymbol{H} = \frac{\partial L}{\partial \boldsymbol{B}} = \frac{1}{\mu}\boldsymbol{B},\quad \boldsymbol{D} = -\frac{\partial L}{\partial \boldsymbol{E}} = \varepsilon\boldsymbol{E};$$

并且由变分法 $\delta\iint L\mathrm{d}\boldsymbol{x}\mathrm{d}t = 0$ 可推导出 Maxwell 电磁场方程组.

为此，通过

$$\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}, \quad \boldsymbol{E} = -\nabla\varphi - \frac{\partial \boldsymbol{A}}{\partial t}$$

引进电磁场的矢势 $\mathbf{A}$ 与标势 φ 作为变分函数. 于是，

$$\begin{aligned}\delta\iint L\mathrm{d}\boldsymbol{x}\mathrm{d}t &= \delta\iint \frac{1}{2}\left(\frac{1}{\mu}B^2 - \varepsilon E^2\right)\mathrm{d}\boldsymbol{x}\mathrm{d}t = \iint\left(\frac{1}{\mu}\boldsymbol{B}\cdot\delta\boldsymbol{B} - \varepsilon\boldsymbol{E}\cdot\delta\boldsymbol{E}\right)\mathrm{d}\boldsymbol{x}\mathrm{d}t \\ &= \iint\left\{\boldsymbol{H}\cdot(\nabla\times\delta\boldsymbol{A}) + \boldsymbol{D}\cdot\left(\nabla(\delta\varphi) + \frac{\partial}{\partial t}(\delta\boldsymbol{A})\right)\right\}\mathrm{d}\boldsymbol{x}\mathrm{d}t.\end{aligned}$$

因为由矢量导数运算公式可知，

$$\boldsymbol{H}\cdot(\nabla\times\delta\boldsymbol{A}) = (\delta\boldsymbol{A})\cdot(\nabla\times\boldsymbol{H}) + \nabla\cdot((\delta\boldsymbol{A})\times\boldsymbol{H}),$$

$$\boldsymbol{D}\cdot\frac{\partial}{\partial t}(\delta\boldsymbol{A}) = -(\delta\boldsymbol{A})\cdot\frac{\partial}{\partial t}\boldsymbol{D} + \frac{\partial}{\partial t}(\boldsymbol{D}\cdot(\delta\boldsymbol{A})),$$

$$\boldsymbol{D}\cdot\nabla(\delta\varphi) = -(\delta\varphi)(\nabla\cdot\boldsymbol{D}) + \nabla\cdot((\delta\varphi)\boldsymbol{D});$$

所以在固定边界的情形下有

$$\delta\iint L\mathrm{d}\boldsymbol{x}\mathrm{d}t = \iint\left\{\left(\nabla\times\boldsymbol{H} - \frac{\partial\boldsymbol{D}}{\partial t}\right)\cdot(\delta\boldsymbol{A}) - (\nabla\cdot\boldsymbol{D})(\delta\varphi)\right\}\mathrm{d}\boldsymbol{x}\mathrm{d}t.$$

最后，由电磁势 $\mathbf{A}$，φ 的定义以及变分法 $\delta\iint L\mathrm{d}\boldsymbol{x}\mathrm{d}t = 0$ 给出

$$\begin{cases}\nabla\cdot\boldsymbol{B} = \nabla\cdot(\nabla\times\boldsymbol{A}) = 0, \\ \nabla\times\boldsymbol{E} + \dfrac{\partial\boldsymbol{B}}{\partial t} = -\nabla\times(\nabla\varphi) = 0, \\ \nabla\cdot\boldsymbol{D} = 0, \\ \nabla\times\boldsymbol{H} - \dfrac{\partial\boldsymbol{D}}{\partial t} = 0;\end{cases}$$

这就是没有电荷与电流情况下的 **Maxwell 方程组.**

存在电荷密度 ρ 与电流密度 $\boldsymbol{j}$ 的情况下，若取 **Lagrange 密度函数**为

$$L = \frac{1}{2}\left(\frac{1}{\mu}B^2 - \varepsilon E^2\right) - \boldsymbol{j}\cdot\mathbf{A} + \rho\varphi,$$

在 ρ，$\boldsymbol{j}$ 不变时给出变分导数

$$[L]_{\varphi(\boldsymbol{x},t)} = -\nabla\cdot\boldsymbol{D} + \rho,$$

$$[L]_{A(\boldsymbol{x},t)} = \nabla\times\boldsymbol{H} - \frac{\partial\boldsymbol{D}}{\partial t} - \boldsymbol{j}.$$

变分法的结果给出 **Maxwell 电磁场方程组**为

$$\begin{cases}\nabla \cdot \boldsymbol{B} = 0, \\ \nabla \times \boldsymbol{E} + \dfrac{\partial \boldsymbol{B}}{\partial t} = 0, \\ \nabla \cdot \boldsymbol{D} = \rho, \\ \nabla \times \boldsymbol{H} - \dfrac{\partial \boldsymbol{D}}{\partial t} = \boldsymbol{j}.\end{cases}$$

9.7.3 Schrödinger 波动力学方程

考虑单个粒子在势场 $V(\boldsymbol{x})$中的定态运动，其 Lagrange 密度函数取为

$$L = -\frac{1}{2}(\nabla\psi^*)\cdot(\nabla\psi) - \psi^* V\psi,$$

这里 ψ 是波函数，ψ^* 是其复共轭，它们还需满足归一化条件

$$g[\psi] = \int G\mathrm{d}\boldsymbol{x} = \int(\psi^*\psi - \delta(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} = 0;$$

为了方便，已令粒子质量 $m=1$ 及 Planck 常量 $\hbar=1$.

注意到

$$(\nabla\psi^*)\cdot(\nabla\psi) = -\psi^*(\nabla^2\psi) + \nabla\cdot(\psi^*\nabla\psi),$$

以及在边界处波函数为零，变分法 $\delta\int(L+EG)\mathrm{d}\boldsymbol{x} = 0$ 给出

$$[L+EG]_{\psi^*(x)} = \left\{\frac{1}{2}\nabla^2 + E - V\right\}\psi = 0,$$

或者将常量写出为

$$\left\{-\frac{\hbar^2}{2m}\nabla^2 + V\right\}\psi = E\psi,$$

这就是单粒子定态运动的 **Schrödinger 方程**；其中作为 Lagrange 乘子引入的 E 是粒子的定态能量. 应该指出，变分导数$[L+Eg]_{\psi(x)}=0$ 给出其共轭方程.

9.7.4 小结

应该指出，变分法在物理学中主要用于：(1) **近似求解**和(2) **归纳定律**.

关于变分法用于近似求解的具体例子可参考作者所著《理论物理基础》[①]一书的第 9.4 节和第 11 章；数值计算时常采用 Ritz 方法(见本书第 9.4 节).

至于变分法用于归纳定律的例子见本书前几小节.

由前面几小节所讨论过的内容可以推知，差不多一切自然定律都能用变分原理的形式予以表达；即，对于不同领域的现象，选择一个合适的 Lagrange 函数，则

① 彭桓武，徐锡申：《理论物理基础》，北京大学出版社，北京，2011.

变分法 $\delta\int L\mathrm{d}t = 0$ 给出该现象规律的描述. 应指出，变分 $\delta\int L\mathrm{d}t$ 本身并不是自然定律，由变分法所给出的 Euler 方程才是自然定律.

为什么物理定律都能用变分原理的形式予以表达呢？这个问题还没有完善的回答. 彭曾向 Schrödinger 问过这个问题，他的回答是，变分方程多是相容的. 这意思是说，可能是因为随便猜一个 Lagrange 函数，由变分原理 $\delta\int L\mathrm{d}t = 0$ 所得到的微分方程组都是相容而不相抵触的缘故. 偏微分方程组的相容性是一个基本问题，只有满足相容条件才可能有解. 所以，由变分法 $\delta\int L\mathrm{d}t = 0$ 给出的 Euler 方程的相容性，这或许是为什么物理定律通常都可写成变分原理形式的主要理由.

第 10 章　微分方程绪论

本书此后三章将转入关于微分方程的讨论.这一章首先要介绍有关基本概念以及几个一般性理论问题.后两章将分别专门讨论二阶线性偏微分方程和常微分方程的解法.

10.1　引　　言

10.1.1　微分方程的有关定义

凡含有自变量、未知函数及其[常]导数或偏导数的方程称为**常微分方程**或**偏微分方程**;而**微分方程**则是二者的总称.

一个微分方程中所含最高阶导数的阶数,以及此导数的幂指数,分别称为**微分方程的阶和次**.

一个微分方程中所出现的未知函数及其各阶导数,若均为一次幂,则称为**线性微分方程**,否则的话,称为**非线性微分方程**.非线性方程中,若对未知函数的最高阶导数来说是线性的,则称为**拟线性微分方程**.

10.1.2　常微分方程示例

(1) 含一个自变量 x 和一个未知函数 $y(x)$ 的情形(下面需要时用 ODE 表示常微分方程,并且用 L,NL,QL 分别表示线性、非线性、拟线性):

$$\frac{\mathrm{d}y}{\mathrm{d}x}=f(x)\qquad\text{(一阶[一次]L)},$$

$$\frac{\mathrm{d}y}{\mathrm{d}x}=f(x,y)\qquad\text{(一阶[一次]L 或 QL)},$$

$$F\left(x,y,\frac{\mathrm{d}y}{\mathrm{d}x}\right)=0\quad\text{(一阶)},$$

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+y^2=0\qquad\text{(二阶[一次]QL)},$$

$$\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2+y^2=0\quad\text{(一阶二次 NL)}.$$

但是，一阶二次 NL ODE

$$\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 - y^2 = 0,$$

实际上可化为两个一阶 L ODE：

$$\frac{\mathrm{d}y}{\mathrm{d}x} - y = 0 \quad 和 \quad \frac{\mathrm{d}y}{\mathrm{d}x} + y = 0,$$

它们是等价的. 另一方面，二阶[一次]QL ODE

$$F\left(x, y\frac{\mathrm{d}y}{\mathrm{d}x}, \frac{\mathrm{d}^2 y}{\mathrm{d}x^2}\right) = \pm\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2} + \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = 0,$$

则与二阶二次 NL ODE

$$1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 - \left(\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}\right)^2 = 0$$

是等价的.

(2) 含一个自变量 x 和二个未知函数 $y_1(x)$，$y_2(x)$ 的情形：因为要确定二个未知函数需二个方程，所以有下述定义：

定义 当方程组中方程个数等于(或小于或大于)未知函数个数时，则该方程组称为**确定**(或**欠定**，或**超定**)**方程组**.

例如，一个方程

$$F\left(x; y_1, y_2; \frac{\mathrm{d}y_1}{\mathrm{d}x}, \frac{\mathrm{d}y_2}{\mathrm{d}x}\right) = 0$$

为欠定方程组，二个方程

$$\begin{cases} F_1\left(x; y_1, y_2; \dfrac{\mathrm{d}y_1}{\mathrm{d}x}, \dfrac{\mathrm{d}y_2}{\mathrm{d}x}\right) = 0, \\ F_2\left(x; y_1, y_2; \dfrac{\mathrm{d}y_1}{\mathrm{d}x}, \dfrac{\mathrm{d}y_2}{\mathrm{d}x}\right) = 0 \end{cases}$$

为确定方程组；三个方程

$$\begin{cases} F_1\left(x; y_1, y_2; \dfrac{\mathrm{d}y_1}{\mathrm{d}x}, \dfrac{\mathrm{d}y_2}{\mathrm{d}x}\right) = 0, \\ F_2\left(x; y_1, y_2; \dfrac{\mathrm{d}y_1}{\mathrm{d}x}, \dfrac{\mathrm{d}y_2}{\mathrm{d}x}\right) = 0, \\ F_3\left(x; y_1, y_2; \dfrac{\mathrm{d}y_1}{\mathrm{d}x}, \dfrac{\mathrm{d}y_2}{\mathrm{d}x}\right) = 0 \end{cases}$$

为超定方程组.

10.1.3 偏微分方程示例

含有两个或两个以上自变量的微分方程属于偏微分方程(简写为 PDE).

例如，对于一个未知函数 $y(x_1, x_2)$ 的情形，一阶偏微分方程

$$A_1\frac{\partial y}{\partial x_1}+A_2\frac{\partial y}{\partial x_2}+By=0$$

中，若系数 A_1,A_2,B 都仅是 x_1,x_2 的函数，则方程是线性的；若 A_1,A_2 中包含 y，则方程是拟线性的. 一阶非线性偏微分方程则一般可表示为

$$F\left(x_1,x_2;y;\frac{\partial y}{\partial x_1},\frac{\partial y}{\partial x_2}\right)=0.$$

对于二阶偏微分方程，亦可类似地予以表述.

对于两个未知函数的情形，

$$F\left(x_1,x_2;y_1,y_2;\frac{\partial y_1}{\partial x_1},\frac{\partial y_1}{\partial x_2},\frac{\partial y_2}{\partial x_1},\frac{\partial y_2}{\partial x_2}\right)=0$$

为欠定[偏微分]方程组；

$$\begin{cases}F_1\left(x_1,x_2;y_1,y_2;\dfrac{\partial y_1}{\partial x_1},\dfrac{\partial y_1}{\partial x_2},\dfrac{\partial y_2}{\partial x_1},\dfrac{\partial y_2}{\partial x_2}\right)=0,\\ F_2\left(x_1,x_2;y_1,y_2;\dfrac{\partial y_1}{\partial x_1},\dfrac{\partial y_1}{\partial x_2},\dfrac{\partial y_2}{\partial x_1},\dfrac{\partial y_2}{\partial x_2}\right)=0\end{cases}$$

为确定[偏微分]方程组；

$$\begin{cases}F_1\left(x_1,x_2;y_1,y_2;\dfrac{\partial y_1}{\partial x_1},\dfrac{\partial y_1}{\partial x_2},\dfrac{\partial y_2}{\partial x_1},\dfrac{\partial y_2}{\partial x_2}\right)=0,\\ F_2\left(x_1,x_2;y_1,y_2;\dfrac{\partial y_1}{\partial x_1},\dfrac{\partial y_1}{\partial x_2},\dfrac{\partial y_2}{\partial x_1},\dfrac{\partial y_2}{\partial x_2}\right)=0,\\ F_3\left(x_1,x_2;y_1,y_2;\dfrac{\partial y_1}{\partial x_1},\dfrac{\partial y_1}{\partial x_2},\dfrac{\partial y_2}{\partial x_1},\dfrac{\partial y_2}{\partial x_2}\right)=0\end{cases}$$

为超定[偏微分]方程组.

10.2 微分方程的等价问题

可以证明，一个高阶微分方程总可以化为一阶微分方程组. 对于常微分方程，反之亦然；而对于偏微分方程，反之却不必然.

10.2.1 常微分方程的等价定理

定理 一个高阶常微分方程与一阶常微分方程组是等价的.

首先以二阶常微分方程

$$F\left(x,y,\frac{\mathrm{d}y}{\mathrm{d}x},\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right)=0$$

为例证来阐明此定理的正确性.

显然，若令 $\frac{\mathrm{d}y}{\mathrm{d}x}=y_1$，则有 $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}y_1}{\mathrm{d}x}$；于是，上述二阶方程化为下列一阶

ODE 组：

$$\begin{cases}\dfrac{\mathrm{d}y}{\mathrm{d}x}-y_1=0,\\ F\left(x;y,y_1;\dfrac{\mathrm{d}y_1}{\mathrm{d}x}\right)=0.\end{cases}$$

另一方面，对于下列一阶常微分方程组

$$\begin{cases}F_1\left(x;y_1,y_2;\dfrac{\mathrm{d}y_1}{\mathrm{d}x},\dfrac{\mathrm{d}y_2}{\mathrm{d}x}\right)=0,\\ F_2\left(x;y_1,y_2;\dfrac{\mathrm{d}y_1}{\mathrm{d}x},\dfrac{\mathrm{d}y_2}{\mathrm{d}x}\right)=0;\end{cases}$$

若将 $y_2,\dfrac{\mathrm{d}y_2}{\mathrm{d}x}$ 当做未知量来求解，根据**隐函数定理**[①]，有解的条件是

$$\frac{\partial(F_1,F_2)}{\partial\left(y_2,\dfrac{\mathrm{d}y_2}{\mathrm{d}x}\right)}=\frac{\partial F_1}{\partial y_2}\frac{\partial F_2}{\partial\dfrac{\mathrm{d}y_2}{\mathrm{d}x}}-\frac{\partial F_1}{\partial\dfrac{\mathrm{d}y_2}{\mathrm{d}x}}\frac{\partial F_2}{\partial y_2}\neq 0.$$

假设此条件是满足的，而对方程组的求解得到

---

① **隐函数**是由 $F(x,y)=0$ 所给出的函数 $y=f(x)$.

隐函数定理可表述如下：如果 $F(x,y)$ 在 (x^0,y^0) 邻域内连续可微，且 $F(x^0,y^0)=0,\left.\dfrac{\partial F}{\partial y}\right|_{(x^0,y^0)}\neq 0$，则有唯一的连续函数 $y=f(x)$ 存在，而且满足 $F(x,f(x))=0,y^0=f(x^0)$.

因为由复合函数微分法

$$\frac{\partial F}{\partial x}\mathrm{d}x+\frac{\partial F}{\partial y}\mathrm{d}y=0 \quad 得 \quad \frac{\mathrm{d}y}{\mathrm{d}x}=-\left(\frac{\partial F}{\partial y}\right)^{-1}\left(\frac{\partial F}{\partial x}\right),$$

由于 $\left.\dfrac{\partial F}{\partial y}\right|_{(x^0,y^0)}\neq 0$，可求出 $\left(\dfrac{\mathrm{d}y}{\mathrm{d}x}\right)$ 以及其高阶导数在 (x^0,y^0) 点之值，即可得 y 的 Taylor 展开，故有 $y=f(x)$ 存在.

对于由 $F_i(x_1,\cdots,x_n;y_1,\cdots,y_m)(i=1,\cdots,m)$ 所给出的隐函数 $y_i=f_i(x_1,\cdots,x_n)(i=1,\cdots,m)$，上述隐函数定理同样成立；只要将 x,y,F,f 视为 $x=(x_1,\cdots,x_n),y=(y_1,\cdots,y_m),F=(F_1,\cdots,F_m),f=(f_1,\cdots,f_m)$，而条件 $\left.\dfrac{\partial F}{\partial x}\right|_{(x^0,y^0)}\neq 0$ 则应换为 Jacobi 行列式 $\left.\dfrac{\partial(F_1,\cdots,F_m)}{\partial(y_1,\cdots,y_m)}\right|_{(x^0,y^0)}\neq 0$.

因为在此情况下，由

$$\frac{\partial F_k}{\partial x_i}+\sum_{j=1}^{m}\frac{\partial F_k}{\partial y_j}\frac{\partial f_j}{\partial x_i}=0 \quad (k=1,\cdots,m;i=1,\cdots,n)$$

中可解出 $\dfrac{\partial f_j}{\partial x_i}$ 的条件是

$$\frac{\partial(F_1,\cdots,F_m)}{\partial(y_1,\cdots,y_m)}\neq 0.$$

其实，前面的证明依然有效. 当然，要将 $\dfrac{\mathrm{d}y}{\mathrm{d}x}$ 换为 $\dfrac{\partial y}{\partial x}$，并注意到应用矩阵运算，$\dfrac{\partial y}{\partial x},\dfrac{\partial F}{\partial x}$ 均为 $m\times n$ 矩阵，$\dfrac{\partial F}{\partial y}$ 为 $m\times m$ 矩阵；而 $\left(\dfrac{\partial F}{\partial y}\right)$ 有逆矩阵 $\left(\dfrac{\partial F}{\partial y}\right)^{-1}$ 的条件是 $\det\left(\dfrac{\partial F}{\partial y}\right)\neq 0$，即 $\dfrac{\partial(F_1,\cdots,F_m)}{\partial(y_1,\cdots,y_m)}\neq 0$.

$$\begin{cases} y_2 = \varphi\left(x; y_1; \dfrac{\mathrm{d}y_1}{\mathrm{d}x}\right), \\ \dfrac{\mathrm{d}y_2}{\mathrm{d}x} = \psi\left(x; y_1; \dfrac{\mathrm{d}y_1}{\mathrm{d}x}\right). \end{cases}$$

消去 y_2，得 $\psi - \dfrac{\mathrm{d}\varphi}{\mathrm{d}x} = 0$，即

$$\psi\left(x; y_1; \frac{\mathrm{d}y_1}{\mathrm{d}x}\right) - \left\{\frac{\partial\varphi}{\partial x} + \frac{\partial\varphi}{\partial y_1}\frac{\mathrm{d}y_1}{\mathrm{d}x} + \frac{\partial\varphi}{\partial \dfrac{\mathrm{d}y_1}{\mathrm{d}x}}\frac{\mathrm{d}^2 y_1}{\mathrm{d}x^2}\right\} = 0;$$

这是一个二阶常微分方程.

从这个例证中可以看出**等价**(用符号$\Longleftrightarrow$表示)的含义：

2 阶 ODE (1个方程) $\Longleftrightarrow$ 1 阶 ODE 组 (2个方程)

这个例证中所采用的方法同样可推广用于对 n 阶情况的证明.

考虑下列 n 阶常微分方程：

$$F\left(x; y; \frac{\mathrm{d}y}{\mathrm{d}x}, \cdots, \frac{\mathrm{d}^n y}{\mathrm{d}x^n}\right) = 0.$$

显然可见，它等价于下列 n 个一阶常微分方程组：

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} - y_1 = 0, \\ \dfrac{\mathrm{d}y_1}{\mathrm{d}x} - y_2 = 0, \\ \quad\vdots \\ \dfrac{\mathrm{d}y_{n-2}}{\mathrm{d}x} - y_{n-1} = 0, \\ F\left(x; y, y_1, \cdots, y_{n-1}; \dfrac{\mathrm{d}y_{n-1}}{\mathrm{d}x}\right) = 0. \end{cases}$$

另一方面，对于下列一阶常微分方程组

$$F_i\left(x; y_1, \cdots, y_n; \frac{\mathrm{d}y_1}{\mathrm{d}x}, \cdots, \frac{\mathrm{d}y_n}{\mathrm{d}x}\right) = 0 \quad (i = 1, 2, \cdots, n);$$

如果满足下列条件：

$$J = \frac{\partial(F_1, \cdots, F_n)}{\partial\left(\dfrac{\mathrm{d}y_1}{\mathrm{d}x}, \cdots, \dfrac{\mathrm{d}y_n}{\mathrm{d}x}\right)} \neq 0,$$

则根据隐函数定理，可解得

$$\frac{\mathrm{d}y}{\mathrm{d}x}=f(x;y,y_1,\cdots,y_{n-1}),$$

$$\frac{\mathrm{d}y_i}{\mathrm{d}x}=f_i(x;y,y_1,\cdots,y_{n-1})\quad(i=1,\cdots,n-1);$$

为了方便，这里已令 $y_n=y$. 现在引进下列算符

$$\theta=\frac{\partial}{\partial x}+f\frac{\partial}{\partial y}+\sum_{r=1}^{n-1}f_r\frac{\partial}{\partial y_r},$$

于是有

$$\frac{\mathrm{d}y}{\mathrm{d}x}=f(x;y,y_1,\cdots,y_{n-1}),$$

$$\frac{\mathrm{d}^m y}{\mathrm{d}x^m}=\theta^{m-1}f(x;y,y_1,\cdots,y_{n-1})\quad(m=2,\cdots,n).$$

从这 n 个方程中消去变量 $y_1,\cdots,y_{n-1}$，即可得到一个 n 阶常微分方程.

以上结果可以简记为

$$\boxed{\underset{(\text{一个方程})}{n\text{ 阶 ODE}}\iff\underset{(n\text{个方程})}{\text{一阶 ODE 组}}}$$

显然，还可推广到高阶常微分方程组的情况：

由 n_1 阶，$\cdots$，n_m 阶方程(共 m 个)构成的高阶常微分方程组等价于 $\sum_{r=1}^{m}n_r$ 个方程构成的一阶常微分方程组；可简记为

$$\boxed{\underset{(m\text{个方程})}{\text{高阶 ODE 组}}\iff\underset{\left(\sum_{r=1}^{m}n_r\text{个方程}\right)}{\text{一阶 ODE 组}}}$$

所以，**常微分方程的最普遍形式就是一阶常微分方程组**.

10.2.2 偏微分方程的等价问题

对于偏微分方程，情况有点不同；上述定理只有一半是对的：

(i) 一个高阶偏微分方程总可化为一阶偏微分方程组；

(ii) 一阶偏微分方程组一般不一定能化至单一未知函数的一个高阶偏微分方程.

首先用一个二阶偏微分方程

$$F(x,y;u;u_x,u_y;u_{xx},u_{xy},u_{yy})=0$$

为例来阐明(i)的正确性. 显然，它与一阶方程组

$$\begin{cases}\dfrac{\partial u}{\partial x}-p=0,\\ \dfrac{\partial u}{\partial y}-q=0,\\ \dfrac{\partial p}{\partial y}-\dfrac{\partial q}{\partial x}=0,\\ F\left(x,y;u,p,q;\dfrac{\partial p}{\partial x},\dfrac{\partial p}{\partial y},\dfrac{\partial q}{\partial y}\right)=0\end{cases}$$

是等价的. 这四个方程中有三个是独立的,正好有三个未知函数,所以是确定方程组.

下面再用两个方程的一阶偏微分方程组

$$\begin{cases}f(x,y;u,v;u_x,v_x,u_y,v_y)=0, & ①\\ g(x,y;u,v;u_x,v_x,u_y,v_y)=0, & ②\end{cases}$$

为例来阐明(ii)的论述.

第一阶段: 这里是两个方程,两个未知函数 $u(x,y)$, $v(x,y)$,要想使之变成一个未知函数 u 的微分方程,需消去三个未知量 v, v_x, v_y,显然是办不到的.

第二阶段: 将方程对 x, y 各求导一次,连原方程共得六个方程,即:

$$①,②,①_x,①_y,②_x,②_y.$$

但由于求导的结果,要消去的未知量亦变为六个. 即

$$v,v_x,v_y;v_{xx},v_{xy},v_{yy},$$

仍然不成.

第三阶段: 再求导一次,得到 12 个方程

$$①,②,①_x,①_y,②_x,②_y,①_{xx},①_{xy},①_{yy},②_{xx},②_{xy},②_{yy};$$

要消去的未知量变为 10 个,即

$$v,v_x,v_y,v_{xx},v_{xy},v_{yy},v_{xxx},v_{xxy},v_{xyy},v_{yyy};$$

所以消去的结果是两个方程.

由此可见,偏数分方程组与高阶偏微分方程不能等价,需要单独研究.

然而,偏微分方程组,包括高阶的,一般来说,永远可以缩小范围当做一阶的.

10.3 初值问题解的存在性定理①

10.3.1 存在性定理

由于一阶偏微分方程组包罗一切微分方程,所以在证明存在性定理时,只须讨

① 参考 10.6 节所引[1]Courant, Hilbert 的书,Band Ⅱ, Kap. Ⅰ, §7;或[8]Goursat 的书. Goursat 书上的证明好一些.

论它即可.

考虑一阶偏微分方程组的**正规形式**[①]

$$\frac{\partial u_i}{\partial x}=f_i\left(x,y;u_1,\cdots,u_m;\frac{\partial u_1}{\partial y},\cdots,\frac{\partial u_m}{\partial y}\right),\quad (i=1,2,\cdots,m)$$

其中 x,y 为自变量,$u_i(x,y)$ $(i=1,\cdots,m)$为未知函数,在给定 $x=0$ 时之初值

$$u_i(0,y)=\varphi_i(y),\quad \left.\frac{\partial u_i}{\partial y}\right|_{x=0}=\varphi_i'(y)$$

的条件下,来讨论方程组之解的存在性,此即所谓**初值问题**.

可以证明,上述初值问题可化为**一阶拟线性偏微分方程组**的下列特定形式

$$\frac{\partial u_i}{\partial x}=\sum_{k=1}^{n}G_{ik}(u_1,\cdots,u_n)\frac{\partial u_k}{\partial y}\quad (i=1,\cdots,n)$$

的等价初值问题,即在 $u_i(0,y)=\varphi_i(y)$这一预给的初值条件下求解 $u_i(x,y)$. 注意,一般有 $n>m$.

如果微分方程 $f_i=\dfrac{\partial u_i}{\partial x}$与初始函数 $\varphi_i(i=1,\cdots,m)$都是解析的,就可求得 u_i 的各阶偏导数在 $x=0,y=0$ 处之值,从而可得 u_i 展开成下列形式幂级数

$$u_i(x,y)=\sum_{k,l}c_{i,kl}x^ky^l,$$

其中系数由微分方程及初始数据唯一确定. 同样,由于初始数据的解析性,$\varphi_i(y)$和 $G_{ik}(u_1,\cdots,u_n)$亦可分别展开成 y 和 u_i 的幂级数. 并且可知,$c_{i,kl}$是后二者展开系数的多项式.

现在可采用**强函数法**[②]来论证上述形式幂级数的收敛性. 构造原来初值问题的强问题,可得假设的幂级数

$$v_i(x,y)=\sum_{k,l}C_{i,kl}x^ky^l,$$

并且可知有

$$C_{i,kl}\geqslant|c_{i,kl}|,$$

即 $v_i(x,y)$也是 $u_i(x,y)$的强函数.

例如,挑选 v_i 使满足(强问题的拟线性方程)

---

① 一个常微分方程,若能将其最高阶导数显式解出,即得此**常微分方程的正规形式**,如

$$\frac{d^ny}{dx^n}=f\left(x;y,\frac{dy}{dx},\cdots,\frac{d^{n-1}y}{dx^{n-1}}\right).$$

对于一个**偏微分方程组**,其**正规形式**为

$$\frac{\partial^k u_i}{\partial x^k}=f_i\left(x,y_1,\cdots,y_l;u_1,\cdots,u_m,\frac{\partial u_1}{\partial x},\cdots,\frac{\partial^{k-1}u_m}{\partial x^{k-1}},\frac{\partial u_1}{\partial y_1},\cdots,\frac{\partial^k u_m}{\partial y_l^k}\right)\quad (i=1,\cdots,m).$$

② 对用幂级数表示的函数,它的**强函数**是将给定级数的系数全用不小于其绝对值的相应系数替换之后的结果.

$$\frac{\partial v_i}{\partial x}=\sum_{k=1}^{n}\frac{M}{1-\dfrac{v_1+\cdots+v_n}{r}}\frac{\partial v_k}{\partial y},$$

$$v_i(0,y)=\frac{My}{\rho-y};$$

其中 ρ 和 r 分别是 $\varphi_i(y)$ 和 $G_{ik}(u_1,\cdots,u_n)$ 的幂级数展开式之收敛半径，而 M 是使强问题中相应展式都满足强函数要求的一个常数.

因为上述方程对 $i=1,\cdots,n$ 均相同，故而有

$$\frac{\partial v}{\partial x}=\frac{nM}{1-\dfrac{nv}{r}}\frac{\partial v}{\partial y},$$

$$v(0,y)=v_0=\frac{My}{\rho-y};$$

由此容易得出关于解的二次方程

$$(v+M)\left[\left(1-\frac{n}{r}v\right)y+nMx\right]=\rho v\left(1-\frac{n}{r}v\right),$$

并且 v 在原点邻域内能展成 x 及 y 的收敛幂级数. 所以，$u_i(x,y)$ 的幂级数是收敛的.

总之，若 f_i 在 $x=0,y=y_0$ 邻域是解析的，而 φ_i 对 y 是解析的，则方程的解存在、唯一，并且解析；解本身满足初值条件. 此即**初值问题的存在性定理**.

由于高阶微分方程均可化为一阶偏微分方程组，而且一般来说，可将之化为正规形式. 由此，所有微分方程，只要解析就有解，当然要有初值条件.

10.3.2　简单例子

下面再用几个简单例子来说明一下以上证明的可行性.

例 1　考虑一阶偏微分方程

$$p=f(x,y,u,q)$$

在给定初始条件 $u(0,y)=\varphi(y)$ 下的初值问题. 这里按照惯例采用 $p=u_x=\dfrac{\partial u}{\partial x}$，$q=u_y=\dfrac{\partial u}{\partial y}$的记号.

若以 u,p,q 作为函数，则得与之相应的一阶拟线性偏微分方程组和初始条件分别为

$$\begin{cases}u_x=p,\\ q_x=p_y,\\ p_x=f_x+f_up+f_qp_y;\end{cases}\qquad\begin{cases}u(0,y)=\varphi(y),\\ q(0,y)=\varphi'(y),\\ p(0,y)=f(0,y,\varphi(y),\varphi'(y)).\end{cases}$$

由于 $p_y=q_x$，即 $u_{xy}=q_x$，积分得

$$u_y(x,y) = q(x,y) + \alpha(y) \to u_y = q,$$

后一步利用了 $u_y(0,y)=q(0,y)$定出 $\alpha(y)=0$. 又因

$$u_{xx} = p_x = \left(\frac{\partial}{\partial x} f(x,y,u,q)\right)_{y\text{不变}},$$

积分后得

$$u_x(x,y) = f(x,y,u,q) + \beta(y) \to p = f,$$

后一步利用了对 p 的初条件定出$\beta(y)=0$. 由此可见,此方程组的解即为原问题的解;也就是说,两者是等价的.

上述拟线性方程组中尚含有自变量 x,y,可采用再引进两个函数 $\xi(x,y)=x$ 和 $\eta(x,y)=y$ 这种技巧,而将之变为系数中不显含 x,y 且对导数为齐次的等价拟线性微分方程组的特殊形式. 对于五个函数 u,p,q,ξ,η 的等价方程组及相应初始条件分别为

$$\begin{cases} u_x = p\eta_y, \\ p_x = f_q p_y + (f_x + f_u p)\eta_y, \\ q_x = p_y, \\ \xi_x = \eta_y, \\ \eta_x = 0; \end{cases} \qquad \begin{cases} u(0,y) = \varphi(y), \\ p(0,y) = f(0,y,\varphi(y),\varphi'(y)), \\ q(0,y) = \varphi'(y), \\ \xi(0,y) = 0, \\ \eta(0,y) = y. \end{cases}$$

应该注意,要将左边方程中 f_x, f_u, f_q 内的 x,y 换为 ξ,η.

例 2　考虑二阶偏微分方程

$$u_{xx} - u_{yy} = 0$$

在给定初始条件 $u(0,y)=\varphi(y)$, $p(0,y)=\psi(y)$下的初值问题.

选定四个函数

$$u(x,y),\ p(x,y) = \frac{\partial u}{\partial x},\ q(x,y) = \frac{\partial u}{\partial y},\ \eta(x,y) = y;$$

于是,等价初值问题的方程组与初条件分别为

$$\begin{cases} u_x = p\eta_y, \\ p_x = q_y, \\ q_x = p_y, \\ \eta_x = 0; \end{cases} \qquad \begin{cases} u(0,y) = \varphi(y), \\ p(0,y) = \psi(y), \\ q(0,y) = \varphi'(y), \\ \eta(0,y) = y. \end{cases}$$

例 3　考虑下列方程

$$u_{xx} - u_y = 0$$

在给定初始条件 $u(0,y)=\varphi(y)$, $p(0,y)=\psi(y)$下的初值问题.

设定三个函数, u,p,η,则等价初值问题为

$$\begin{cases} u_x = p\eta_y, \\ p_x = u_y, \\ \eta_x = 0; \end{cases} \quad \begin{cases} u(0,y) = \varphi(y), \\ p(0,y) = \psi(y), \\ \eta(0,y) = y. \end{cases}$$

10.3.3 通解中的任意常数或任意函数

常(或偏)微分方程的通解必需包含任意常数(或函数).

例如,对 n 阶常微分方程的正规形式

$$\frac{\mathrm{d}^n y}{\mathrm{d}x^n} = F\left(x, y, \frac{\mathrm{d}y}{\mathrm{d}x}, \frac{\mathrm{d}^2 y}{\mathrm{d}x^2}, \cdots, \frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}}\right),$$

给定在 $x=x_0$ 时的初值条件:

$$y(x_0) = c_0, \left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=x_0} = c_1, \cdots, \left.\frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}}\right|_{x=x_0} = c_{n-1};$$

若微分方程 F 是解析的,则其解在 $x=x_0$ 邻域可展开成 Taylor 级数

$$y(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \left.\frac{\mathrm{d}^k y}{\mathrm{d}x^k}\right|_{x=x_0} (x-x_0)^k = f(x, c_0, c_1, \cdots, c_{n-1});$$

显然,其中从 $k=0$ 至 $k=n-1$ 的系数 c_k 由初值条件直接给出,而 $k=n$ 及其后的系数则由微分方程及对之一再求导间接给出为 $c_0, c_1, \cdots, c_{n-1}$ 之函数. 如

$$\left.\frac{\mathrm{d}^n y}{\mathrm{d}x^n}\right|_{x=x_0} = F(x_0, c_0, c_1, \cdots, c_{n-1}) \equiv F_0,$$

$$\left.\frac{\mathrm{d}^{n+1} y}{\mathrm{d}x^{n+1}}\right|_{x=x_0} = \left(\frac{\partial}{\partial x_0} + c_1 \frac{\partial}{\partial c_0} + \cdots + c_{n-1} \frac{\partial}{\partial c_{n-2}} + F_0 \frac{\partial}{\partial c_{n-1}}\right) F_0,$$

$$\vdots$$

$$\left.\frac{\mathrm{d}^{n+l} y}{\mathrm{d}x^{n+l}}\right|_{x=x_0} = \left(\frac{\partial}{\partial x_0} + c_1 \frac{\partial}{\partial c_0} + \cdots + c_{n-1} \frac{\partial}{\partial c_{n-2}} + F_0 \frac{\partial}{\partial c_{n-1}}\right)^l F_0.$$

由此可见,n 阶常微分方程的通解中包含 n 个任意常数.

对于一阶拟线性偏微分方程组的初值问题

$$\frac{\partial u_i}{\partial x} = \sum_{k=1}^{m} G_{ik} \frac{\partial u_k}{\partial y},$$

$$u_i(0,y) = \varphi_i(y), \quad \left.\frac{\partial u_k}{\partial y}\right|_{x=0} = \varphi_k'(y);$$

涉及 m 个函数 $\varphi_i(y)$ $(i=1,\cdots,m)$. 特别是对 $m=1$, $\frac{\partial u}{\partial x} = G(u)\frac{\partial u}{\partial y}$ 涉及一个任意函数 $\varphi(y)$.

10.3.4 微分方程与差分方程

还可以采用另一种看法. 就 y 而言,任意函数 $\varphi(y)$ 可看作无穷个常数的集合.

粗略地说，可用差分来代替微分，

$$\left(\frac{\partial u}{\partial y}\right)_{y=y_j}=\frac{u(x,y_{j+1})-u(x,y_j)}{\Delta y}.$$

若记 $u(x,y_j)=u_j(x)$，则偏微分方程$\frac{\partial u}{\partial x}=G(u)\frac{\partial u}{\partial y}$可变成

$$\frac{\mathrm{d}u_j}{\mathrm{d}x}=g(u_j)(u_{j+1}-u_j)\quad(j=0,1,2,\cdots),$$

其中 $g(u_j)=\frac{1}{\Delta y}G(u_j)$. 由此可见，一个偏微分方程，其复杂的程度就等于一组无穷个常微分方程.

再用这种看法，可把所有的微分方程变成**差分方程**. 若记$u(x_i,y_j)=u_{i,j}$，则上述偏微分方程$\frac{\partial u}{\partial x}=G(x)\frac{\partial u}{\partial y}$进一步还可变成①

$$\frac{u_{i+1,j}-u_{i,j}}{\Delta x}=G(u_{i,j})\frac{u_{i,j+1}-u_{i,j}}{\Delta y}\quad(i,j=0,1,2,\cdots).$$

也就是说，一个含二自变量的偏微分方程就可当做二重无穷代数方程组去解.

10.4　边值问题解的方法

关于微分方程边值问题解的存在性和唯一性问题，情况比较复杂，没有像初值问题那样普遍的存在性定理. 因此这里不作专门讨论，而将在以后对某些类型方程的求解中有所论述. 不过应指出，就物理上的实际问题而言，一定有解，只是所给定边界值是否相容. **如果边界条件是相容的和解析的，则解一定是存在的、唯一的和解析的**.

对于具体求解方法，或者是直接解边值问题，或者是先求通解，然后凑足边界值.

下面举几个常微分方程的例子.

例 1

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+2\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2=0,$$

$$y(0)=1,\quad y(1)=\frac{1}{2}.$$

---

①　这里用的是**向前差分格式**，类似地有**向后差分格式**，统称**单侧格式**. 还有常用的是**中心差分格式**，上述微分方程变成

$$\frac{u_{i+1,j}-u_{i-1,j}}{2\Delta x}=G(u_{i,j})\frac{u_{i,j+1}-u_{i,j-1}}{2\Delta y};$$

当然，对端点处需另行处理.

令 $p(x)=\frac{\mathrm{d}y}{\mathrm{d}x}$,得$\frac{\mathrm{d}p}{\mathrm{d}x}=-2p^2$,解出

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{1}{2(x+a)},$$

再积分得 $y=\frac{1}{2}\ln(x+a)+b$. 由边界条件定出 a,b 后得到方程的解为

$$y(x)=1+\frac{1}{2}\ln\left(1+\frac{1-\mathrm{e}}{\mathrm{e}}x\right).$$

注意,当方程中缺未知量 y 时总可以先解降一阶的方程,再积分一次即得原方程之解.

例 2
$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+y=1,$$
$$y(0)=0,\quad y(1)=0.$$

由于此线性微分方程的解在两端点为零,所以可用积分变换方法来解.

考虑函数 $y(x)$的 Fourier 级数

$$y(x)=\sum_{n=1}^{\infty}b_n\sin\left(\frac{n\pi x}{l}\right),$$
$$b_n=\frac{2}{l}\int_0^l y(x)\sin\left(\frac{n\pi x}{l}\right)\mathrm{d}x.$$

于是,令 $l=1$,用 $2\int_0^1\cdots\sin(n\pi x)\mathrm{d}x$ 作用于方程后得到

$$2\int_0^1\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\sin(n\pi x)\mathrm{d}x+2\int_0^1 y(x)\sin(n\pi x)\mathrm{d}x=2\int_0^1\sin(n\pi x)\mathrm{d}x.$$

将左边第一项进行分部积分二次,并利用边界条件,而右边可直接积出,上式变为

$$(1-n^2\pi^2)\cdot 2\int_0^1 y(x)\sin(n\pi x)\mathrm{d}x=\frac{2(1-\cos n\pi)}{n\pi},$$
$$b_n=\frac{2(1-\cos n\pi)}{n\pi(1-n^2\pi^2)}.$$

注意到 $\cos n\pi=(-1)^n$,结果求得微分方程之解为

$$\begin{aligned}y(x)&=\sum_{n=1}^{\infty}\frac{2(1-(-1)^n)}{n\pi(1-n^2\pi^2)}\sin(n\pi x)\\&=\sum_{k=0}^{\infty}\frac{4}{(2k+1)\pi[1-(2k+1)^2\pi^2]}\sin[(2k+1)\pi x].\end{aligned}$$

例 3 考虑二阶线性非齐次常微分方程的边值问题

$$\frac{\mathrm{d}^2u}{\mathrm{d}z^2}+p(z)\frac{\mathrm{d}u}{\mathrm{d}z}+q(z)u=f(z),$$
$$u(0)=a,\quad u(1)=b;$$

可以先求解两个相关的初值问题

$$\frac{\mathrm{d}^2 v}{\mathrm{d}z^2} + p(z)\frac{\mathrm{d}v}{\mathrm{d}z} + q(z)v = f(z),$$

$$v(0) = a, \quad \left.\frac{\mathrm{d}v}{\mathrm{d}z}\right|_{z=0} = 1;$$

$$\frac{\mathrm{d}^2 w}{\mathrm{d}z^2} + p(z)\frac{\mathrm{d}w}{\mathrm{d}z} + q(z)w = f(z),$$

$$w(0) = a, \quad \left.\frac{\mathrm{d}w}{\mathrm{d}z}\right|_{z=0} = 0.$$

于是,原方程的边值问题之解可表达为两初值问题之解的叠加

$$u(z) = \frac{1}{w(1) - v(1)}\left[(b - v(1))w(z) + (w(1) - b)v(z)\right].$$

10.5 一阶偏微分方程的一般理论[①]

一个一阶偏微分方程,其普遍理论与力学中的 Hamilton-Jacobi 理论[②]是等价的.

在关于变分法的一章(9.7.1 节)中曾经论证了 Hamilton-Jacobi 偏微分方程

$$\frac{\partial G}{\partial t} + H\left(t, q_i, \frac{\partial G}{\partial q_i}\right) = 0$$

与 Hamilton 正则方程

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \quad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i}$$

之间的等价.

现在来论证一下,任何一个一阶偏微分方程都可变成 **Hamilton-Jacobi 偏微分方程**那样的形式.

如果未知函数 u 不显现于微分方程中,例如

$$F\left(x_1, x_2, \cdots, x_{n+1}, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \cdots, \frac{\partial u}{\partial x_{n+1}}\right) = 0,$$

若其中某个导数,如$\frac{\partial u}{\partial x_{n+1}}$可解出. 现在令 $x_{n+1} = x$,则方程变为

$$\frac{\partial u}{\partial x} + H\left(x, x_1, \cdots, x_n, \frac{\partial u}{\partial x_1}, \cdots, \frac{\partial u}{\partial x_n}\right) = 0;$$

此即 Hamilton-Jacobi 偏微分方程的形式(注意到,x, x_i, u 分别相当于 t, q_i, G).

---

① 参考 10.6 节所引[1]Courant, Hilbert 的书, Bd. Ⅱ, Kap. Ⅱ, §8, 2—3.

② 关于力学中的 Hamiltom-Jacobi 理论可参考:彭桓武,徐锡申:《理论物理基础》,北京大学出版社,北京,2011,1.4 节,7—10 页.

如果微分方程中显含未知函数 u,例如,

$$F\left(x_1,\cdots,x_n,u,\frac{\partial u}{\partial x_1},\cdots,\frac{\partial u}{\partial x_n}\right)=0,$$

可引入 $u=x_{n+1}$ 作为附加自变量,并将解族表示为隐式

$$u=u(x_1,\cdots,x_n,c)\to\varphi(x_1,\cdots,x_n,x_{n+1},c)=0.$$

现在若将 $\frac{\partial u}{\partial x_i}=-\frac{\partial\varphi}{\partial x_i}\Big/\frac{\partial\varphi}{\partial x_{n+1}}$ 代入原方程,即得到关于新未知函数 φ 的一阶导数的微分方程

$$\widetilde{F}\left(x_1,\cdots,x_n,x_{n+1},\frac{\partial\varphi}{\partial x_1},\cdots,\frac{\partial\varphi}{\partial x_n},\frac{\partial\varphi}{\partial x_{n+1}}\right)=0,$$

其中不再显含 φ 函数本身.于是,又可按上述办法将之化为 Hamilton-Jacobi 方程的形式.

应指出,一个二阶或高阶偏微分方程,无如上之普遍理论.

10.6 微分方程参考书

下面列出关于微分方程的一些参考书.

[1] R. Courant, D. Hilbert, *Methoden der Mathematisch Physik*, Bd. Ⅱ, Springer, Berlin, 1924.[中译本,R 柯朗,D. 希尔伯特《数理物理方法》,卷Ⅱ,熊振翔、杨易辰译,科学出版社,北京,1977.]

[2] Bieberbach, *Theorie der Gewöhnlichen Differentialgleichungen*, Berlin, 1953.

[3] A. Sommerfeld, *Partial Differential Equations in Physics*, Academic Press, New York, 1949.

[4] H. Bateman, *Partial Differential Equations of Mathematical Physics*, Cambridge, New York, 1944.

[5] A. R. Forsyth, *A Treatise on Differential Equations*, Cambridge, New York, 1914.

[6] A. R. Forsyth, *Theory of Diffential Equations*, Vols Ⅰ-Ⅵ, Cambridge, New York, 1906.

Port Ⅰ, Vol. Ⅰ, *Exact Equations and Pfaff's Problem* (1890).

Part Ⅱ, Vols. Ⅱ, Ⅲ, *Ordinary Equations, Not Linear* (1899).

Part Ⅲ, Vol. Ⅳ, *Ordinary Linear Equations* (1902).

Part Ⅳ, Vols. Ⅴ, Ⅵ, *Partial Differential Equations* (1906).

[7] E. L. Ince, *Ordinary Differential Equations*, Longmans, New York, 1927. 这是英文书中较好的.

[8] E. Goursat, *Cours d'Analyse*, Paris, Ⅰ(1910), Ⅱ(1911), Ⅲ(1915).

[9] E. T. Whittaker, G. N. Watson, *A Course of Modern Analysis*, Cambridge University Press, Cambridge, 4th ed., 1940.

还可参考以下书籍.

[10] 王竹溪、郭敦仁,《特殊函数概论》,北京大学出版社,北京,2000.(此系新版,初版由科学出版社于 1965 年出版.)

[11] 郭敦仁,《数学物理方法》,人民教育出版社,北京,1965.

[12] D. Zwillinger, *Handbook of Differential Equations*, Academic, Boston, 1989.

[13] P. M. Morse, H. Feshbach, *Methods of Theoretical Physics*, McGraw-Hill, New York, 1953.

第 11 章　二阶线性偏微分方程

11.1　分类和举例

本章将专门论述**二阶线性偏微分方程**.

对于两个自变量 x,y 的情况,其一般形式为

$$A\frac{\partial^2 u}{\partial x^2}+2B\frac{\partial^2 u}{\partial x\partial y}+C\frac{\partial^2 u}{\partial y^2}+D\frac{\partial u}{\partial x}+E\frac{\partial u}{\partial y}+Fu=0,$$

其中系数 A,B,C,D,E,F 都仅是 x,y 的函数,不再包含 u. 若 A,B,C 中包含 u,则方程是**拟线性的**. 同时注意到,对于**线性方程**,u 的一阶和二阶导数都必需是线性的,否则就成为**非线性方程**.

根据**判别式** $AC-B^2$ 的符号,可将上述二阶线性偏微分方程划分为下列三种类型:

椭圆型方程　$AC-B^2>0$,

双曲型方程　$AC-B^2<0$,

抛物型方程　$AC-B^2=0$.

对于含 n 个自变量 $x_1,\cdots,x_n$ 的情况,一般形式为

$$\sum_{i,j=1}^{n}a_{ij}\frac{\partial^2 u}{\partial x_i\partial x_j}+\sum_{i=1}^{n}b_i\frac{\partial u}{\partial x_i}+cu=0,$$

其中系数 $a_{ij}=a_{ji},b_i,c$ 等都仅是 $x_1,\cdots,x_n$ 的函数. 对此可按以下方式进行分类.

在 u 的定义域上每一点,由系数矩阵 $\{a_{ij}\}$ 构成的二次型 $Q=\sum\limits_{i,j}a_{ij}\xi_i\xi_j$ 经主轴变换后可化为 $Q=\sum\limits_{i}\lambda_i\zeta_i^2$,其中 $\{\lambda_i\}$ 为矩阵 $\{a_{ij}\}$ 的本征值. 于是可根据本征值组 $\{\lambda_i\}$ 来对上述偏微分方程进行分类.

(i) 若 $\{\lambda_i\}$ 均不为零且同号(同为正或同为负),则方程为**椭圆型的**.

(ii) 若 $\{\lambda_i\}$ 均不为零,且有一个与其余为异号,则方程为**双曲型的**.

(iii) 若 $\{\lambda_i\}$ 中有任何一个为零,则方程为**抛物型的**.

应该指出,二阶偏微分方程的类型完全由其二阶导数的系数矩阵 $\{a_{ij}\}$ 所确定,因而上述分类方式对下列形式的方程:

$$\sum_{i,j=1}^{n}a_{ij}(x_1,\cdots,x_n)\frac{\partial^2 u}{\partial x_i\partial x_j}+f\left(x_1,\cdots,x_n;u;\frac{\partial u}{\partial x_1},\cdots,\frac{\partial u}{\partial x_n}\right)=0,$$

亦适用.

下面举出这三种类型偏微分方程的典型例子,并阐述其物理含义.

(1) 势方程

椭圆型方程的典型例子是 **Laplace 方程**:

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad AC - B^2 \equiv 1 > 0.$$

这类型方程属于势论的范畴,亦常称之为**势方程**. 由于空间中一点的势总是某值而不改变,所以,势是属于平衡现象. 因此,椭圆型微分方程在物理上描述一种**平衡现象**.

(2) 波动方程

双曲型方程的典型例子是**波动方程**:

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2};$$

其中特例是

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0, \quad (\text{平面波})$$

$$\frac{\partial^2 u}{\partial r^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} + \frac{2}{r}\frac{\partial u}{\partial r} = 0, \quad (\text{球面波})$$

后二者均有 $AC-B^2=-\frac{1}{c^2}<0$.

这类型方程在物理上描述扰动的传播过程,例如电磁波、声波、流体力学波等的传播过程,属于不平衡现象. 因此,双曲型方程在物理上描述**不平衡现象**.

(3) 扩散方程

抛物型方程的典型例子是**热传导方程(或扩散方程)**:

$$k\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial t} = 0,$$

由于 $A=k, B=0, C=0$,有 $AC-B^2=0$;这里 u 代表温度(热传导)或浓度(扩散).

对于热传导现象,u 代表温度,起先有温度差,后来经传导作用而渐趋消灭. 对于扩散现象,u 代表浓度,起先有浓度差,后来经扩散作用而渐趋消灭. 因此,抛物型微分方程在物理上描述**趋向平衡的现象**.

应该指出,当微分方程内有虚数($\mathrm{i}=\sqrt{-1}$)时,不能用上面的物理解释. 例如,

$$\frac{\partial^2 \psi}{\partial x^2} - \mathrm{i}\frac{\partial \psi}{\partial t} = 0$$

属于抛物型,然而,它在物理上所描述的却是波动这样一种不平衡现象.

11.2 抛物型微分方程的解

11.2.1 热传导方程的物理推导

当仅考虑热传导现象时，有 **Fourier 定律**

$$\boldsymbol{J} = -\kappa \nabla T,$$

其中 $\boldsymbol{J}$ 是热流（单位时间内通过温度梯度方向单位面积所传递的能量），T 是温度，κ 是热导率；另一方面，有能量方程

$$\frac{\partial \varepsilon}{\partial t} + \nabla \cdot \boldsymbol{J} = 0,$$

其中 ε 是单位体积的能量. 若 $\varepsilon = cT$，其中 c 是物体单位体积的热容，则容易求得

$$\kappa \nabla^2 T - c\frac{\partial T}{\partial t} = 0;$$

这就是**热传导方程**. 或者，令 $k = \frac{\kappa}{c}$，$u = T$，故有

$$k\nabla^2 u - \frac{\partial u}{\partial t} = 0.$$

对于纯扩散过程，可类似地进行推导，注意到扩散是物质输运过程，物质流与质量密度 ρ（或浓度 c）之梯度成正比（**Fick 定律**），以及还有连续性方程；结果得到与上述完全相同的方程. 不过此时有 $u = \rho$ 或 c 和 $k = D$，后者为扩散系数.

11.2.2 热传导方程的一般解法

当 $k = \text{const}$ 时，可令 $\tau = kt$，则

$$k\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial t} = 0 \Longrightarrow \frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial \tau} = 0.$$

对于这类常系数线性偏微分方程，都有解，可采用叠加法来求解；即，如果 $u_1, u_2, \cdots, u_n$ 均为其特解，则 $\sum c_i u_i$ 亦为其解. 但应指出，要有无穷个特解才能代替一个任意函数.

Fourier 为了解决此问题，发明了 Fourier 级数，以及 Fourier 积分.

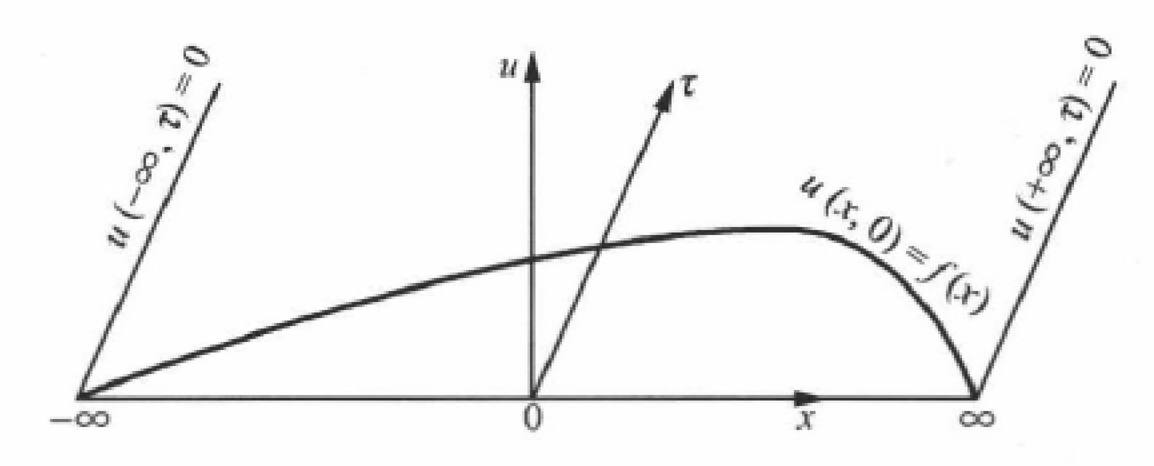
为了求解热传导（或扩散）方程，需要给出初始条件（初始温度或浓度的分布），以及给出边界条件（用以规定，例如，物质能否穿越边界）.

现在来求解

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial \tau} = 0,$$

给定初值和边值分别为（参考下页图）

$$u(x,0)=f(x),\quad -\infty<x<+\infty;$$
$$u(\pm\infty,\tau)=0,\quad \tau>0.$$



根据 Fourier 积分定理(见第 8 章 8.4.2 节)将 $f(x)$用积分表示.由

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(\omega)\mathrm{e}^{-\mathrm{i}\omega x}\mathrm{d}\omega,$$

$$F(\omega)=\int_{-\infty}^{\infty}f(\xi)\mathrm{e}^{\mathrm{i}\omega\xi}\mathrm{d}\xi,$$

得到

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}\omega\int_{-\infty}^{\infty}f(\xi)\mathrm{e}^{-\mathrm{i}\omega(x-\xi)}\mathrm{d}\xi.$$

注意到,若 $u=A\mathrm{e}^{ax+b\tau}$是$\frac{\partial^2 u}{\partial x^2}-\frac{\partial u}{\partial\tau}=0$ 的解,则要求 $b=a^2$;因此,$\mathrm{e}^{-\mathrm{i}\omega(x-\xi)-\omega^2\tau}$是方程的一个解,其中 ω,ξ 为任意常数.

由此可见,无穷个特解 $\mathrm{e}^{-\mathrm{i}\omega(x-\xi)-\omega^2\tau}$的叠加亦为其解.因而,当 $\tau>0$ 时,满足初条件的解是

$$\begin{aligned}u(x,\tau)&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}\omega\int_{-\infty}^{\infty}f(\xi)\mathrm{e}^{-\mathrm{i}\omega(x-\xi)-\omega^2\tau}\mathrm{d}\xi\\&=\int_{-\infty}^{\infty}f(\xi)\mathrm{d}\xi\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}\omega(x-\xi)-\omega^2\tau}\mathrm{d}\omega\right\}\\&=\int_{-\infty}^{\infty}f(\xi)\mathrm{d}\xi\left\{\frac{1}{\sqrt{4\pi\tau}}\mathrm{e}^{-(x-\xi)^2/4\tau}\right\};\end{aligned}$$

这里应用了下列数学公式(见 8.4.6 节之例 2)

$$\int_{-\infty}^{\infty}\mathrm{e}^{-\gamma\zeta^2+\mathrm{i}\alpha\zeta}\mathrm{d}\zeta=\sqrt{\frac{\pi}{\gamma}}\mathrm{e}^{-\alpha^2/4\gamma}.$$

同时注意到,当 $x=\pm\infty$时有 $u(\pm\infty,\tau)=0$;以及当 $\tau\to 0$ 时,{　}内函数趋于 $\delta\{x-\xi\}$而有 $u(x,0)=f(x)$.所以,前述结果的确是所求扩散方程之解.另外,若初条件为 $\tau=\tau_0\,(kt=kt_0)$时 $u(x,\tau_0)=f(x)$,则前式中的 τ 应换成$\tau-\tau_0$.或者,令

$$U(x-\xi,t-t_0)=\frac{1}{\sqrt{4\pi k(t-t_0)}}\mathrm{e}^{-\frac{(x-\xi)^2}{4k(t-t_0)}},$$

则扩散方程的解可写成

$$u(x,t)=\int_{-\infty}^{\infty} f(\xi)\mathrm{d}\xi \frac{1}{\sqrt{4\pi k(t-t_0)}}\mathrm{e}^{-\frac{(x-\xi)^2}{4k(t-t_0)}}$$

$$=\int_{-\infty}^{\infty} f(\xi)\mathrm{d}\xi \cdot U(x-\xi,t-t_0).$$

推广到三维空间，扩散方程形式是

$$\frac{\partial u}{\partial t}=k\nabla^2 u=k\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)u,$$

初条件为，当 $t=t_0$ 时有 $u=f(x,y,z)$；其解是

$$u(x,y,z,t)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi,\eta,\zeta)\mathrm{d}\xi\mathrm{d}\eta\mathrm{d}\zeta\cdot\left\{[4\pi k(t-t_0)]^{-3/2}\right.$$

$$\left.\cdot\exp\left[-\frac{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}{4k(t-t_0)}\right]\right\},$$

这是从长时间看，为**扩散看法**；或者写成

$$u(x,y,z,t)=\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{d}\omega_1\mathrm{d}\omega_2\mathrm{d}\omega_3$$

$$\cdot\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi,\eta,\zeta)\mathrm{d}\xi\mathrm{d}\eta\mathrm{d}\zeta$$

$$\times\exp\{-\mathrm{i}\omega_1(x-\xi)-\mathrm{i}\omega_2(y-\eta)$$

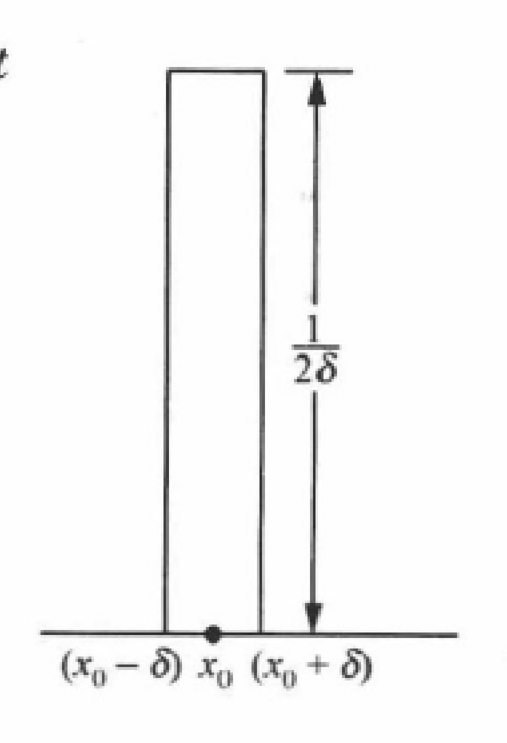
$$-\mathrm{i}\omega_3(z-\zeta)-\omega^2 k(t-t_0)\},$$

其中 $\omega^2=\omega_1^2+\omega_2^2+\omega_3^2$，这是从短时间看，为**波动看法**.

11.2.3 具体例子

(1) 单位热源的时空热量分布

设在 $t=t_0$ 时在 $x=x_0$ 处给定一单位热源，观察其在 $t=t$ 时热量分布的情形. 也就是说，给定初值为(见右图)



$$f(x)=\begin{cases}\frac{1}{2\delta}, & x_0-\delta<x<x_0+\delta,\\ 0, & x<x_0-\delta,\ x_0+\delta<x.\end{cases}$$

于是，在 $t=t$ 时有

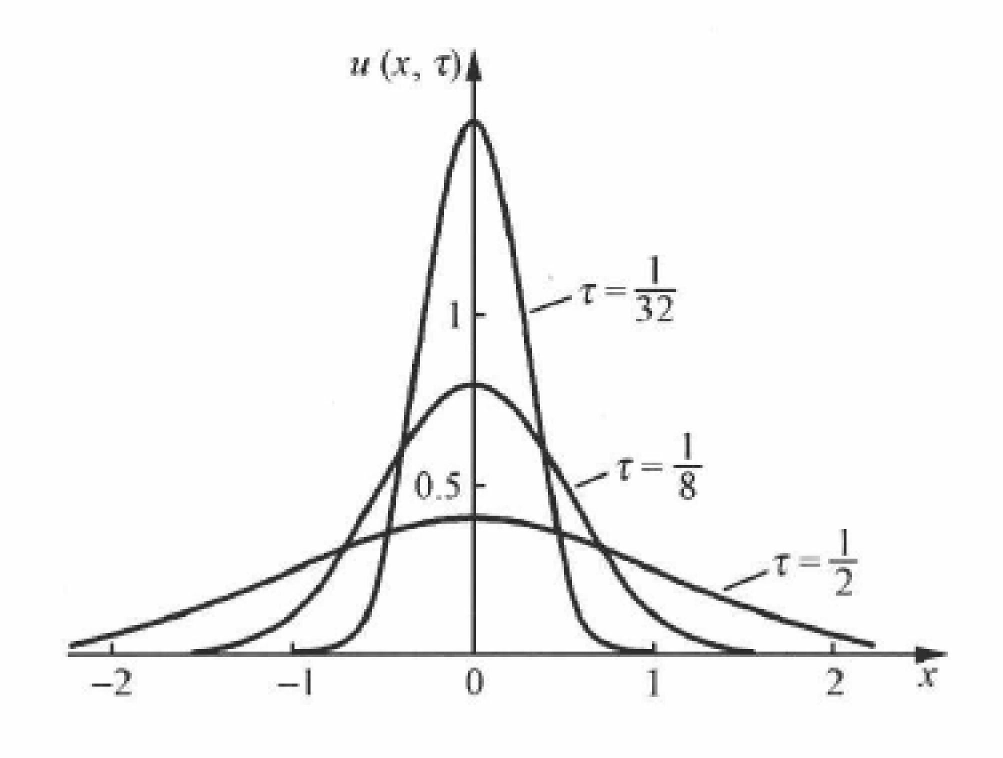
$$u(x,t)=\int_{-\infty}^{\infty} f(\xi)\mathrm{d}\xi U(x-\xi,t-t_0)$$

$$=\int_{x_0-\delta}^{x_0+\delta}\frac{1}{2\delta}\mathrm{d}\xi U(x-\xi,t-t_0)$$

$$=U(x-x_0,t-t_0)$$

$$=\frac{1}{\sqrt{4\pi k(t-t_0)}}\exp\left[-\frac{(x-x_0)^2}{4k(t-t_0)}\right],$$

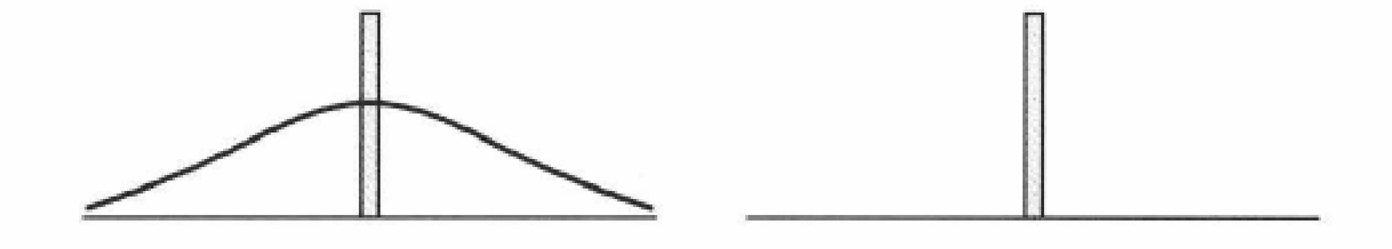
这里应用了中值定理. 若令 $x_0=0, k(t-t_0)=\tau$，则如下图所示.



在任何时候，热量均按 Gauss 分布. 并且，

$$\text{总热量}=\int_{-\infty}^{\infty} U\mathrm{d}x=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{4k(t-t_0)}}\exp\left[-\frac{(x-x_0)^2}{4k(t-t_0)}\right]\mathrm{d}x=1.$$

与单位热源扩散传播相类似之问题，如分子的扩散，可用实验验证此结果. 例如，在水中用两块玻璃板隔出一薄层，在此薄层中注入墨汁(如下右图)；然后迅速抽去玻璃板，观察墨汁扩散的情况(某一时刻如下左图曲线所示).



(2) Brown 运动

Brown 运动(指花粉或胶粒在水中的随机运动)满足扩散方程，一个微粒在时刻 t 处于 x 点的概率为

$$U(x,t)=\frac{1}{\sqrt{4\pi kt}}\mathrm{e}^{-\frac{x^2}{4kt}},$$

$$\int_{-\infty}^{\infty} U(x,t)\mathrm{d}x=1.$$

若令 $x\to x-x_0=\Delta x, t\to t-t_0=\Delta t, a=\dfrac{1}{4kt}$，则有

$$\overline{(\Delta x)^2}=\frac{\int_{-\infty}^{\infty} x^2U\mathrm{d}x}{\int_{-\infty}^{\infty} U\mathrm{d}x}=\int_{-\infty}^{\infty} x^2U\mathrm{d}x$$

$$=\sqrt{\frac{a}{\pi}}\int_{-\infty}^{\infty} x^2\mathrm{e}^{-ax^2}\mathrm{d}x=\sqrt{\frac{a}{\pi}}\left(-\frac{\mathrm{d}}{\mathrm{d}a}\right)\int_{-\infty}^{\infty}\mathrm{e}^{-ax^2}\mathrm{d}x$$

$$=\sqrt{\frac{a}{\pi}}\left(-\frac{\mathrm{d}}{\mathrm{d}a}\right)\sqrt{\frac{\pi}{a}}=\frac{1}{2a}=2kt.$$

实际上,Brown 粒子是在三维空间运动,满足三维扩散方程

$$\frac{\partial u}{\partial t} = k\,\nabla^2 u,$$

其解为

$$u(r,t) = \frac{1}{(4\pi kt)^{3/2}}\mathrm{e}^{-\frac{r^2}{4kt}},$$

其中 $r^2 = x^2 + y^2 + z^2$,故

$$\begin{aligned}
\overline{r^2} &= 4\pi\int_0^\infty r^2 u r^2\,\mathrm{d}r = 4\pi\left(\frac{a}{\pi}\right)^{3/2}\int_0^\infty r^4\mathrm{e}^{-ar^2}\,\mathrm{d}r \\
&= 4\pi\left(\frac{a}{\pi}\right)^{3/2}\left(-\frac{\mathrm{d}}{\mathrm{d}a}\right)^2\int_0^\infty \mathrm{e}^{-ar^2}\,\mathrm{d}r \\
&= 4\pi\left(\frac{a}{\pi}\right)^{3/2}\left(-\frac{\mathrm{d}}{\mathrm{d}a}\right)^2\left[\frac{1}{2}\left(\frac{\pi}{a}\right)^{1/2}\right] \\
&= 4\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{3}{2}\cdot a^{3/2}a^{-5/2} \\
&= \frac{3}{2a} = 6kt.
\end{aligned}$$

由于$\overline{x^2}=\overline{y^2}=\overline{z^2}$,所以有

$$\overline{x^2} = \frac{1}{3}\,\overline{r^2} = 2kt;$$

这是 Einstein 关于 Brown 运动的理论结果.

11.2.4 边界条件与 Green 函数

前面在求解热传导方程(或扩散方程)

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2} \quad 或 \quad \frac{\partial u}{\partial t} = k\,\nabla^2 u$$

时,给出的初值条件是 $t=0$ 时 $u=f(x)$,边值条件是 $x=\pm\infty$时对一切 $t>0$ 有 $u=0$;求得的解是

$$u(x,t) = \int_{-\infty}^{\infty} f(\xi)\,\mathrm{d}\xi\, U(x-\xi,t),$$

$$U(x-\xi,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\omega^2 kt - \mathrm{i}\omega(x-\xi)}\,\mathrm{d}\omega = \frac{1}{\sqrt{4\pi kt}}\mathrm{e}^{-\frac{(x-\xi)^2}{4kt}},$$

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}\omega\int_{-\infty}^{\infty} f(\xi)\,\mathrm{e}^{-\mathrm{i}\omega(x-\xi)}\,\mathrm{d}\xi.$$

这个解称为**主解(基本解)**,相当于没有物理边界的情况.

根据物理现象所给出的实际边界条件总是相容的,就热传导现象而言,具体的边值条件可有下列三种类型:

(i) **等温边界**，例如，在 $x=0$ 和 $x=\infty$ 处温度不变，比如可取为 $u=0$.

(ii) **绝热边界**，在边界上没有热量流动，比如可取 $\dfrac{\partial u}{\partial x}=0$.

(iii) **向外传热的边界**，设边界外温度为零，由边界表面向外辐射出的热流为

$$J=\kappa h u,$$

注意到 $J=-\kappa\dfrac{\partial u}{\partial x}$，即有边界条件

$$\frac{\partial u}{\partial x}+hu=0,$$

这里 κh 称为**表面传热系数**.

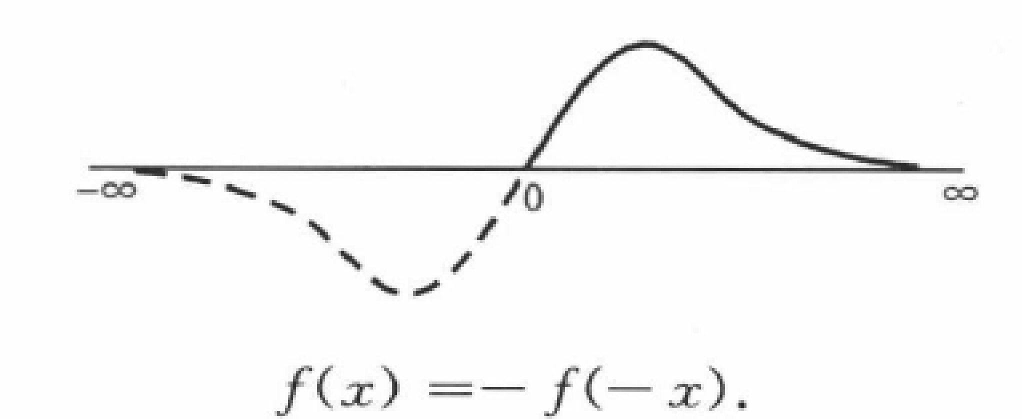
下面再对前二种边界条件之解阐述如下.

(1) 等温边界

初值　$u=f(x)$，$t=0$，$0<x<\infty$；

边值　$u=0$，$x=0$ 和 $x=\infty$，$0<t$.

对于 $t=0$，$-\infty<x<0$，扩展 $f(x)$ 使满足(见下图)



$$f(x)=-f(-x).$$

于是，新的初值和边值变为

初值　$u=f(x)$，$t=0$，$-\infty<x<+\infty$；

边值　$u=0$，($x=0$ 和)$x=\pm\infty$，$0<t$.

这就与前面求解时的条件一致.

因此，其解为

$$\begin{aligned}u(x,t)&=\int_{-\infty}^{\infty}f(\xi)\mathrm{d}\xi U(x-\xi,t)\\&=\int_{0}^{\infty}f(\xi)\mathrm{d}\xi U(x-\xi,t)+\int_{-\infty}^{0}f(\xi)\mathrm{d}\xi U(x-\xi,t),\end{aligned}$$

在后一积分中作变量变换 $\xi\to-\xi$，并且注意到 $f(-\xi)=-f(\xi)$，即有

$$\begin{aligned}\int_{-\infty}^{0}f(\xi)\mathrm{d}\xi U(x-\xi,t)&=-\int_{\infty}^{0}f(-\xi)\mathrm{d}\xi U(x+\xi,t)\\&=-\int_{0}^{\infty}f(\xi)\mathrm{d}\xi U(x+\xi,t).\end{aligned}$$

所以，最后得到

$$u(x,t)=\int_0^\infty f(\xi)\mathrm{d}\xi[U(x-\xi,t)-U(x+\xi,t)]$$

$$=\int_0^\infty f(\xi)\mathrm{d}\xi G_a(x,\xi;t),$$

$$G_a(x,\xi;t)=U(x-\xi,t)-U(x+\xi,t),$$

其中 $G_a(x,\xi;t)$称为此边界条件下的 **Green 函数**.

(2) 绝热边界

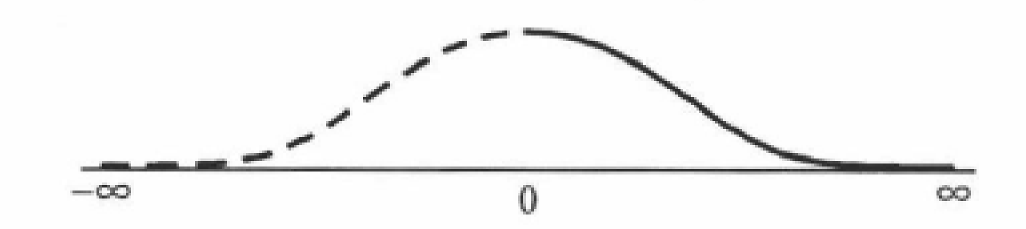
初值 $t=0$ 时，$u=f(x)$，$0<x<\infty$；

边值 $\dfrac{\partial u}{\partial x}=0\ (x=0)$，$u=0\ (x=\infty)$.

对于 $t=0$，将 $f(x)$扩展使(见下图)

$$f(x)=f(-x),\quad -\infty<x<0,$$

这就与前面求解时的条件一致.

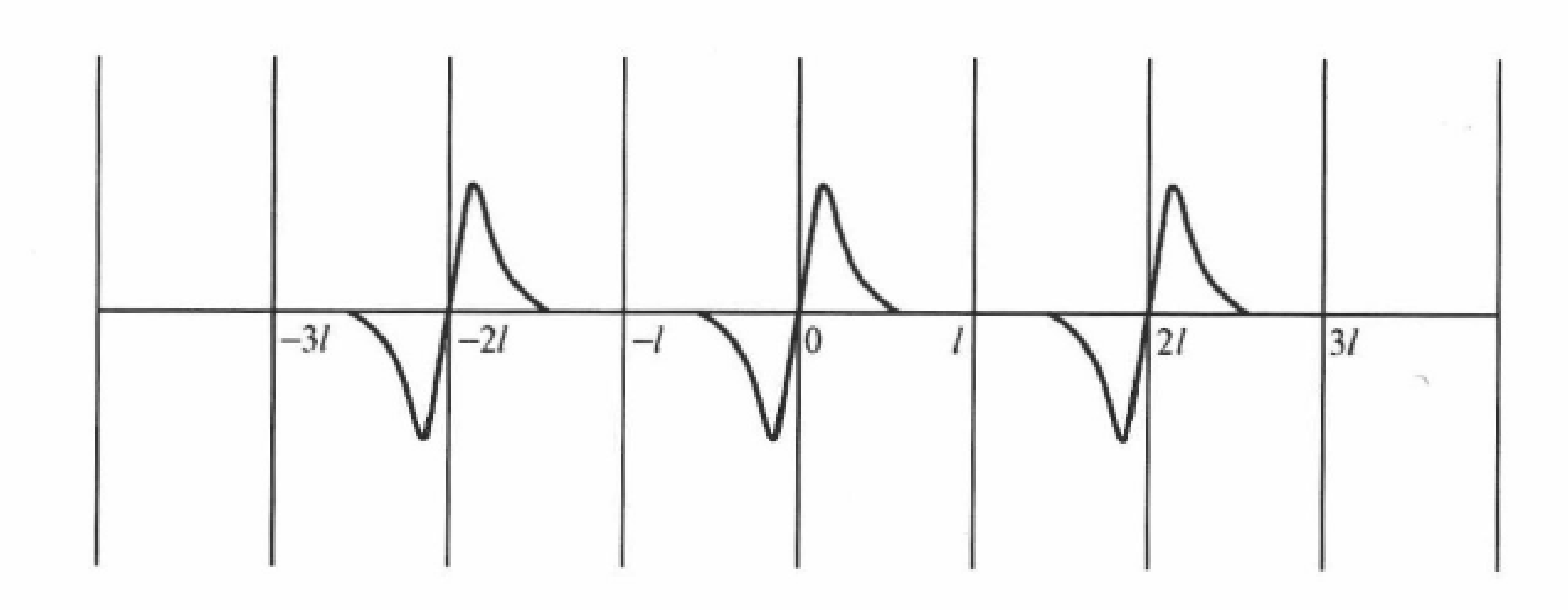


因此，容易求得其解为

$$u(x,t)=\int_0^\infty f(\xi)\mathrm{d}\xi G_b(x,\xi;t),$$

$$G_b(x,\xi;t)=U(x-\xi,t)+U(x+\xi,t).$$

对于**有限长度的边界**，例如，在 $0<x<l$ 内置一单位热源的情况，常采用**镜像法**；如下图所示.



因此，容易求得

$$G(x,\xi;t)=\sum_{n=-\infty}^{\infty}[U(x-\xi-n\cdot 2l;t)-U(x+\xi-n\cdot 2l;t)].$$

11.2.5　地层的热传导问题

设地面($x=0$)温度随时间 t 的平均变化为 $u(0,t)=f(t)$，求地层中温度的平均变化 $u(x,t)$，其中 x 为离地面的深度.

热传导方程是

$$k\frac{\partial^2 u}{\partial x^2}-\frac{\partial u}{\partial t}=0,$$

常量 k 为地层的平均**扩散传热率**.

可以想象到地下极深处($x=\infty$)的温度不会随地面温度而变化，换言之，$u(\infty,t)=c$，c 为常量，为方便起见，可令 $c=0$.

因此，边界条件是，对于 $0<t<T$，

$$u(0,t)=f(t)\quad 和\quad u(\infty,t)=0;$$

这里 T 可以是一年或一日，也可以是其他感兴趣的时间周期.

这个问题可采用 Fourier 级数来解，与前面用 Fourier 积分进行求解有所不同.

将 $f(t)$展成 Fourier 级数为

$$f(t)=\sum_{n=-\infty}^{\infty}c_n\mathrm{e}^{2\pi int/T},$$

若假定所求之解可写成

$$u(x,t)=\sum_{n=-\infty}^{\infty}c_nu_n(x)\mathrm{e}^{2\pi int/T},$$

则有

$$\frac{\partial u}{\partial t}=\sum c_nu_n(x)\frac{2\pi\mathrm{i}n}{T}\mathrm{e}^{2\pi int/T},$$

$$k\frac{\partial^2 u}{\partial x^2}=\sum c_nk\frac{\mathrm{d}^2u_n(x)}{\mathrm{d}x^2}\mathrm{e}^{2\pi int/T}.$$

由热传导方程 $k\frac{\partial^2 u}{\partial x^2}=\frac{\partial u}{\partial t}$要求 $\mathrm{e}^{2\pi int/T}$各项之系数应满足二阶常微分方程

$$\frac{\mathrm{d}^2u_n(x)}{\mathrm{d}x^2}=p_n^2u_n(x),$$

其中 $p_n^2=\frac{2\pi\mathrm{i}n}{kT}$，或者

$$\begin{aligned}p_n&=\sqrt{\frac{\pi}{kT}}\sqrt{2\mathrm{i}n}\\&=\sqrt{\frac{\pi|n|}{kT}}(1\pm\mathrm{i}),\end{aligned}$$

当 $n>0$ 时取"+"号，$n<0$ 时取"−"号.

上述常微分方程之解为

$$u_n(x)=A_n e^{p_n x}+B_n e^{-p_n x};$$

其中系数 A_n 和 B_n 由边界条件确定：

$$x=\infty \text{ 时 } u=0 \Longrightarrow u_n=0 \Longrightarrow A_n=0,$$
$$x=0 \text{ 时 } u=f(t) \Longrightarrow u_n=1 \Longrightarrow B_n=1.$$

因此，令 $q_n=\sqrt{\frac{\pi|n|}{kT}}$，则

$$u_{|n|}(x)=\exp\{-q_n(1+\mathrm{i})x\},$$
$$u_{-|n|}(x)=\exp\{-q_n(1-\mathrm{i})x\}.$$

注意到 $c_n^*=c_{-n}, c_{-n}^*=c_n$，若令

$$c_n=|c_n|e^{\mathrm{i}\gamma_n},$$

则有 $c_{-n}=|c_n|e^{-\mathrm{i}\gamma_n}$. 于是，最后求得

$$\begin{aligned}u(x,t)&=\sum_{n=-\infty}^{\infty}c_n u_n(x)e^{2\pi \mathrm{i}nt/T}\\&=c_0+\sum_{n=1}^{\infty}|c_n|e^{\mathrm{i}\gamma_n}e^{-q_n(1+\mathrm{i})x}e^{2\pi \mathrm{i}nt/T}+\sum_{n=1}^{\infty}|c_n|e^{-\mathrm{i}\gamma_n}e^{-q_n(1-\mathrm{i})x}e^{-2\pi \mathrm{i}nt/T}\\&=c_0+\sum_{n=1}^{\infty}|c_n|e^{-q_n x}\left[e^{\mathrm{i}\left(2\pi n\frac{t}{T}+\gamma_n-q_n x\right)}+e^{-\mathrm{i}\left(2\pi n\frac{t}{T}+\gamma_n-q_n x\right)}\right]\\&=c_0+2\sum_{n=1}^{\infty}|c_n|e^{-q_n x}\cos\left(2\pi n\frac{t}{T}+\gamma_n-q_n x\right).\end{aligned}$$

因为 $q_n=\sqrt{\frac{\pi|n|}{kT}}>0$，所以，这是**阻尼谐振解**.

若令 $q_n\lambda_n=2\pi$，则每当 x 增加 λ_n 时，该项的振幅减小至原来的 $e^{-2\pi}\approx 2\times10^{-3}$.

对平均土壤来说，$k=2\times10^{-7}\ \mathrm{m^2\cdot s^{-1}}$，如果选取 $T=1$ a(年) $=3.15\times10^7$ s，$x=1$ m，则有 $q_1 x\approx 0.71\approx\frac{\pi}{4}$，以及 $e^{-q_1 x}\approx\frac{1}{2}$. 所以，在一米深处就减小到原来的一半.

同时，在 $x=4$ m 处有 $q_1 x\approx\pi$，与表面为异相，即相位滞后半个周期. 所以，当地面是夏天时，地下该处是冬天；而当地面是冬天时，地下该处是夏天. 并且，在地下 4 m 处的振幅已减小到原来的 $\frac{1}{16}$.

另一方面，如果选取 $T=1$ d(日)，则 q_n 变大 $\sqrt{365}\approx 19$ 倍，结果使地面上的温度变化更不容易传到地下去.

11.3　双曲型及椭圆型微分方程的解

11.3.1　引言

(1) D'Alembert 方程

下列二阶线性偏微分方程

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}=-f(x,y,z,t),$$

称为 **D'Alembert 方程**. 若引用三维空间的 **Laplace 符** ∇^2（有时写成 Δ）或四维时空的 **D'Alembert 符** $\Box^2$（有时仅写成 $\Box$），上述方程还可写成

$$\nabla^2 u-\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}=-f,\quad \nabla^2\equiv\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2};$$

$$\Box^2 u=-f,\quad \Box^2\equiv\nabla^2-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}.$$

很显然，由上述方程的系数矩阵所构成的二次型 $Q=\sum_{ij}a_{ij}\xi_i\xi_j=\sum_i\lambda_i\zeta_i^2$，即

$$Q=\xi_x^2+\xi_y^2+\xi_z^2-\frac{1}{c^2}\xi_t^2,$$

其中 $\lambda_x=\lambda_y=\lambda_z=1,\lambda_t=-\frac{1}{c^2}$ 均不为零，且有 λ_t 与其余为异号，所以，微分方程属于双曲型（见 11.1 节）. 这是**非齐次波动方程**.

(2) 椭圆型可看作双曲型的特例

按物理的观点，一切椭圆型微分方程均可看作是双曲型微分方程的特例.

例如，D'Alembert 方程中，当 $\frac{\partial u}{\partial t}=0$ 时或当 u 对 t 为周期函数时，都将化为椭圆型方程. 现在予以阐述如下.

① 当 $\frac{\partial u}{\partial t}=0$ 时，D'Alembert 方程变成 **Poisson 方程**：

$$\nabla^2 u=-f;$$

很显然，这是椭圆型微分方程.

对于 $\frac{\partial u}{\partial t}=0$ 可有两种情形：**静态情形**和**定态情形**.

例如，静止电荷所产生的电场可由电势描述，这是静电现象；又如，恒定电流所产生的磁场可由磁矢势描述，这是静磁现象. 前者属于静态情形，而后者属于定态

情形；因为后一情形下，虽然电流不变，但是存在电荷的恒定流动.

静电势 $\varphi(x,y,z)$ 的 Poisson 方程是

$$\nabla^2\varphi = -\frac{1}{\varepsilon}\rho,$$

其中 $\rho(x,y,z)$ 是已知的电荷密度，ε 是介质的电容率. 注意到静电场强度 $\boldsymbol{E}(x,y,z)$ 由 $\boldsymbol{E}=-\nabla\varphi$ 确定.

静磁现象中的磁矢势 $\boldsymbol{A}(x,y,z)$ 所应满足的方程（矢量式 Poisson 方程）是

$$\nabla^2\boldsymbol{A} = -\mu\boldsymbol{j},$$

其中 $\boldsymbol{j}(x,y,z)$ 是已知的电流密度，μ 是介质的磁导率. 注意到磁感应强度 $\boldsymbol{B}(x,y,z)$ 由 $\boldsymbol{B}=\nabla\times\boldsymbol{A}$ 确定.

还应指出，在区域内无源的情况下，Poisson 方程化为 **Laplace 方程**

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2} = 0.$$

② 对于周期情形，

$$u \propto \mathrm{e}^{\pm \mathrm{i}\omega t},\quad \frac{\partial u}{\partial t} = \pm\,\mathrm{i}\omega u,\quad \frac{\partial^2 u}{\partial t^2} = -\,\omega^2 u,$$

于是，D'Alembert 方程化为 **Helmholtz 方程**

$$\nabla^2 u + k^2 u = -\,f,$$

其中 $k=\frac{\omega}{c}$. 很显然，这是椭圆型微分方程.

就物理学而言，椭圆型很少是四维时空的，也就是说，不涉及时间的变化. 因此，按物理的观点，椭圆型可看作双曲型的特例.

下面要对 D'Alembert 方程进行求解. 首先将给出 Kirchhoff 公式，并讨论一些特殊情形. 接着将给出该公式的证明，最后将采用分离变量法及叠加原理对波动方程进行求解.

11.3.2　D'Alembert 方程的 Kirchhoff 公式①

现在要求解 D'Alembert 方程

$$\Box^2 u = \left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)u = -\,f,$$

其中 $u=u(x,y,z,t)$，$f=f(x,y,z,t)$.

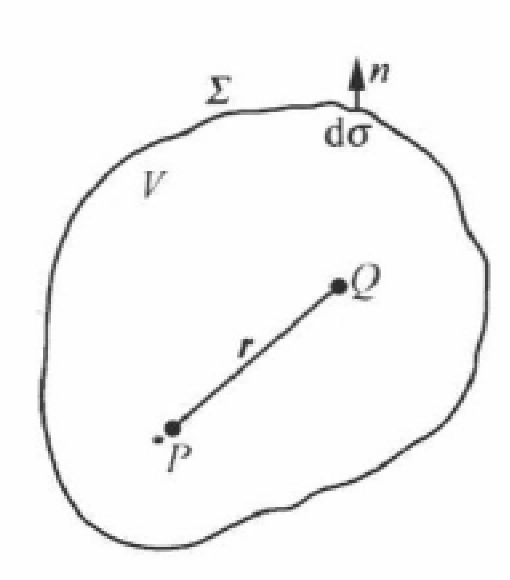
这一类问题多半是电学上的，例如，f 描述电荷（或电流）分布，u 代表电势（或

① 可参考彭桓武、徐锡申：《理论物理基础》第二章习题 2.12，2.26，2.27.

磁矢势).

首先给出 D'Alembert 方程的 **Kirchhoff 公式**：

$$u_Q=\frac{1}{4\pi}\int_V\frac{[f]}{\bar{r}}\mathrm{d}\tau-\frac{1}{4\pi}\int_\Sigma\left\{[u]\frac{\partial}{\partial n}\left(\frac{1}{\bar{r}}\right)-\frac{1}{\bar{r}}\left[\frac{\partial u}{\partial n}\right]-\frac{1}{c\bar{r}}\frac{\partial\bar{r}}{\partial n}\left[\frac{\partial u}{\partial t}\right]\right\}\mathrm{d}\sigma,$$



其中 $\bar{r}$ 是由给定点 Q：(x_1,y_1,z_1) 至区域 V 内一点 P：(x,y,z) 之间的距离(见图)，$\bar{r}^2=(x-x_1)^2+(y-y_1)^2+(z-z_1)^2$，或 $\bar{\boldsymbol{r}}=\boldsymbol{r}-\boldsymbol{r}_1$；$\mathrm{d}\tau=\mathrm{d}x\mathrm{d}y\mathrm{d}z$ 是区域 V 内 P 点处的体积元，$\mathrm{d}\sigma$ 是 V 的外边界 Σ 上某点的面元；$\frac{\partial}{\partial n}$ 指沿该面元外法线方向求导；而[ ]表示取括号内函数的**推迟值**，例如，

$$[f(x,y,z,t)]=f\left(x,y,z,t-\frac{\bar{r}}{c}\right).$$

应该指出，仅当 Q 点在 V 内时才有 $u_Q=u(x_1,y_1,z_1,t)$ 不为零，而当 Q 点在 V 外时则有 $u_Q\equiv 0$.

Kirchhoff 公式本身是个恒等式.具体形式则取决于边界条件和波源的情况.下面讨论物理上常碰到的两类情况.

(1) 无界情形

当源的分布在有限区域以内时，可将边界 Σ 取在无穷远处，显然边界条件应取为在 Σ 上的 $[u]$，$\left[\frac{\partial u}{\partial n}\right]$，$\left[\frac{\partial u}{\partial t}\right]$ 都等于零，因为离源太远处不受源的影响.于是得到

$$u_Q=\frac{1}{4\pi}\int_V\frac{[f]}{\bar{r}}\mathrm{d}\tau.$$

特别对于电学而言，此情况下，

$$\nabla^2\varphi-\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2}=-\frac{\rho}{\varepsilon}$$

的解是**推迟势**

$$\varphi(x_1,y_1,z_1,t)=\frac{1}{4\pi\varepsilon}\int_V\frac{\rho\left(x,y,z,t-\frac{\bar{r}}{c}\right)}{\bar{r}}\mathrm{d}\tau,$$

其中 $\bar{r}=[(x-x_1)^2+(y-y_1)^2+(z-z_1)^2]^{1/2}$，$\mathrm{d}\tau=\mathrm{d}x\mathrm{d}y\mathrm{d}z$.

另一方面，对于静电现象，Poisson 方程

$$\nabla^2\varphi=-\frac{\rho}{\varepsilon}$$

的解是**静电势**

$$\varphi(x_1,y_1,z_1)=\frac{1}{4\pi\varepsilon}\int_V\frac{\rho(x,y,z)}{\bar{r}}\mathrm{d}\tau.$$

(2) 无源情形

对于无源情形，$f=0$，波动方程

$$\nabla^2 u-\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}=0$$

的 Kirchhoff 公式为

$$u_Q=-\frac{1}{4\pi}\int_\Sigma\left\{[u]\frac{\partial}{\partial n}\left(\frac{1}{\bar{r}}\right)-\frac{1}{\bar{r}}\left[\frac{\partial u}{\partial n}\right]-\frac{1}{c\bar{r}}\frac{\partial\bar{r}}{\partial n}\left[\frac{\partial u}{\partial t}\right]\right\}\mathrm{d}\sigma,$$

这就是 **Huygens 原理**的数学表述.

对于**周期情形**，可令

$$u(x,y,z,t)=A(x,y,z)\mathrm{e}^{-\mathrm{i}\omega t},$$

将$\frac{\partial}{\partial t}\to-\mathrm{i}\omega$，波动方程变为 **Helmholtz 方程**，

$$\nabla^2 u-\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}=0\to\nabla^2 A+k^2 A=0,$$

其中 $k=\frac{\omega}{c}$. 因此，只要在 Kirchhoff 公式中作下列变换

$$\frac{\partial}{\partial t}\to-\mathrm{i}\omega,\quad [\mathrm{e}^{-\mathrm{i}\omega t}]\to\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{e}^{\mathrm{i}k\bar{r}},$$

就立即得到 **Helmholtz 公式**

$$A_Q=-\frac{1}{4\pi}\int_\Sigma\left\{A\mathrm{e}^{\mathrm{i}k\bar{r}}\frac{\partial}{\partial n}\left(\frac{1}{\bar{r}}\right)-\frac{\mathrm{e}^{\mathrm{i}k\bar{r}}}{\bar{r}}\frac{\partial A}{\partial n}+A\frac{\mathrm{i}k\mathrm{e}^{\mathrm{i}k\bar{r}}}{\bar{r}}\frac{\partial\bar{r}}{\partial n}\right\}\mathrm{d}\sigma$$

$$=-\frac{1}{4\pi}\int_\Sigma\left\{A\frac{\partial}{\partial n}\left(\frac{\mathrm{e}^{\mathrm{i}k\bar{r}}}{\bar{r}}\right)-\frac{\mathrm{e}^{\mathrm{i}k\bar{r}}}{\bar{r}}\frac{\partial A}{\partial n}\right\}\mathrm{d}\sigma,$$

这是关于 Helmholtz 方程之解的一种特殊推导.

另一方面，对于**静态情形**，可令$\frac{\partial}{\partial t}\to 0$，波动方程变为 **Laplace 方程**：$\nabla^2 v=0$. 只要令 $\omega=0$，$k=0$，就可由 Kirchhoff 公式或 Helmholtz 公式立即得到 Laplace 方程之解

$$v_Q=\frac{1}{4\pi}\int_\Sigma\left\{\frac{1}{\bar{r}}\frac{\partial v}{\partial n}-v\frac{\partial}{\partial n}\left(\frac{1}{\bar{r}}\right)\right\}\mathrm{d}\sigma.$$

Laplace 方程之解称为**球谐函数**，亦称**调和函数**. 由于 v 是环绕 Q 点之球面上平均 v 值，故 v_Q 不能是最大值或最小值.

11.3.3 Kirchhoff 公式的证明

对于线性微分方程，可通过引进伴随微分方程和利用广义 Green 定理来求解. 下面用此方法来给出 Kirchhoff 公式的证明.

(1) 伴随微分方程

对于线性(常或偏)微分方程,可用**线性微分算符** $L(\cdot)$表述为

$$L(u)=0.$$

例如,二阶偏微分算符

$$L(\cdot)=A\frac{\partial^2}{\partial x^2}+2B\frac{\partial^2}{\partial x\partial y}+C\frac{\partial^2}{\partial y^2}+D\frac{\partial}{\partial x}+E\frac{\partial}{\partial y}+F.$$

若有另一线性微分算符 $M(\cdot)$能使

$$vL(u)-uM(v)=\text{完全散度},$$

则 $M(\cdot)$称为 $L(\cdot)$的**伴随微分算符**,而 $M(v)=0$ 称为 $L(u)=0$ 的**伴随微分方程**.

现举例具体说明如下.

(i) 给 $L(u)=\frac{\mathrm{d}u}{\mathrm{d}x}$,求 $M(u)$.

$$vL(u)=v\frac{\mathrm{d}u}{\mathrm{d}x}=\frac{\mathrm{d}}{\mathrm{d}x}(vu)-u\frac{\mathrm{d}v}{\mathrm{d}x},$$

$$vL(u)+u\frac{\mathrm{d}v}{\mathrm{d}x}=\frac{\mathrm{d}}{\mathrm{d}x}(vu)=\text{完全散度};$$

$$M(v)=-\frac{\mathrm{d}v}{\mathrm{d}x}.$$

或者说,$\frac{\mathrm{d}}{\mathrm{d}x}$的伴随算符是$-\frac{\mathrm{d}}{\mathrm{d}x}$.

(ii) 给 $L(u)=\nabla^2u$,求 $M(v)$.

因为 $L(u)=\nabla^2u=\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)u$,而

$$v\frac{\partial^2}{\partial x^2}u=\frac{\partial}{\partial x}\left(v\frac{\partial u}{\partial x}\right)-\frac{\partial v}{\partial x}\frac{\partial u}{\partial x}=\frac{\partial}{\partial x}\left(v\frac{\partial u}{\partial x}-u\frac{\partial v}{\partial x}\right)+u\frac{\partial^2}{\partial x^2}v,$$

容易求得

$$v\nabla^2u-u\nabla^2v=\nabla\cdot(v\nabla u-u\nabla v),$$

这是 **Green 定理的微分形式**. 由此可见,

$$M(v)=\nabla^2v,$$

即有 $M=L$,L 称为**自伴微分算符**.

(iii) 对于二阶线性偏微分算符 $L(\cdot)$,

$$L(u)=A\frac{\partial^2u}{\partial x^2}+2B\frac{\partial^2u}{\partial x\partial y}+C\frac{\partial^2u}{\partial y^2}+D\frac{\partial u}{\partial x}+E\frac{\partial u}{\partial y}+Fu,$$

求 $M(\cdot)$.

因为

$$vA\frac{\partial^2 u}{\partial x^2}=\frac{\partial}{\partial x}\left(vA\frac{\partial u}{\partial x}\right)-\frac{\partial u}{\partial x}\frac{\partial}{\partial x}(vA)$$

$$=\frac{\partial}{\partial x}\left(vA\frac{\partial u}{\partial x}-u\frac{\partial}{\partial x}(vA)\right)+u\frac{\partial^2}{\partial x^2}(vA),$$

$$vB\frac{\partial^2 u}{\partial x\partial y}=\frac{\partial}{\partial x}\left(vB\frac{\partial u}{\partial y}\right)-\frac{\partial}{\partial y}\left(u\frac{\partial}{\partial x}(vB)\right)+u\frac{\partial^2}{\partial x\partial y}(vB)$$

$$=\frac{\partial}{\partial y}\left(vB\frac{\partial u}{\partial x}\right)-\frac{\partial}{\partial x}\left(u\frac{\partial}{\partial y}(vB)\right)+u\frac{\partial^2}{\partial x\partial y}(vB),$$

$$vD\frac{\partial u}{\partial x}=\frac{\partial}{\partial x}(vDu)-u\frac{\partial}{\partial x}(vD),$$

等等，所以有

$$vL(u)-uM(v)=\frac{\partial}{\partial x}\left\{\left(vA\frac{\partial u}{\partial x}-u\frac{\partial}{\partial x}(vA)\right)+\left(vB\frac{\partial u}{\partial y}-u\frac{\partial}{\partial y}(vB)\right)+vDu\right\}$$

$$+\frac{\partial}{\partial y}\left\{\left(vC\frac{\partial u}{\partial y}-u\frac{\partial}{\partial y}(vC)\right)+\left(vB\frac{\partial u}{\partial x}-u\frac{\partial}{\partial x}(vB)\right)+vEu\right\},$$

$$M(v)=\frac{\partial^2}{\partial x^2}(vA)+2\frac{\partial^2}{\partial x\partial y}(vB)+\frac{\partial^2}{\partial y^2}(vC)$$

$$-\frac{\partial}{\partial x}(vD)-\frac{\partial}{\partial y}(vE)+(vF).$$

由此可见，求出微分算符之伴随算符的**简捷方法**是：将 L 中每一项的次序颠倒之，再将$\frac{\partial}{\partial x}$等变为$\left(-\frac{\partial}{\partial x}\right)$等即得 M. 如

L	$A\left(\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial x}\right)$	$B\left(\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial y}\right)$	$D\left(\frac{\partial}{\partial x}\right)$	F
M	$\left(-\frac{\partial}{\partial x}\right)\left(-\frac{\partial}{\partial x}\right)A$	$\left(-\frac{\partial}{\partial y}\right)\left(-\frac{\partial}{\partial x}\right)B$	$\left(-\frac{\partial}{\partial x}\right)D$	F

另一方面，对于常系数线性微分方程，且仅含偶阶导数项时，显然有 $M=L$(**自伴算符**)，称为**自伴线性微分方程**.

(2) 广义 Green 定理

前面已经给出 Green 定理的微分形式

$$u\nabla^2 v-v\nabla^2 u=\nabla\cdot(u\nabla v-v\nabla u),$$

容易得出 **Green 定理的积分形式为**

$$\int_V\{u\nabla^2 v-v\nabla^2 u\}\mathrm{d}\tau=\int_\Sigma\{u\nabla v-v\nabla u\}\cdot\boldsymbol{n}\mathrm{d}\sigma,$$

其中 $\mathrm{d}\tau=\mathrm{d}x\mathrm{d}y\mathrm{d}z$ 是体积 V 中的体元，$\mathrm{d}\sigma$ 是包围 V 的边界面积 Σ 上的面元，而 $\boldsymbol{n}$ 是此面元的外法线方向的单位矢量.

对于一般线性微分算符 $L(\cdot)$，可类似地给出**广义 Green 定理**，其微分形式和

积分形式可分别表达为

$$vL(u)-uM(v)=\operatorname{div}\boldsymbol{P}(u,v),$$

$$\int_V\{vL(u)-uM(v)\}\mathrm{d}\tau=\int_\Sigma \boldsymbol{P}(u,v)\cdot\boldsymbol{n}\mathrm{d}\sigma;$$

这里 div 表示广义散度.

(3) Kirchhoff 公式的具体证明①

为了证明 Kirchhoff 公式,可令

$$v(\boldsymbol{r},t)=u\left(\boldsymbol{r},t-\frac{\bar{r}}{c}\right)=[u(\boldsymbol{r},t)],$$

其中 $\bar{r}=|\boldsymbol{r}-\boldsymbol{r}_1|$. 首先来求$\nabla v$,

$$\begin{aligned}\nabla v&=\nabla u\left(\boldsymbol{r},t-\frac{\bar{r}}{c}\right)=\nabla[u(\boldsymbol{r},t)]\\&=\left[\left(\nabla-(\nabla\bar{r})\frac{\partial}{c\partial t}\right)u\right]=[\nabla u]-(\nabla\bar{r})\left[\frac{\partial u}{c\partial t}\right].\end{aligned}$$

可见,[ ]前的算符∇,移至[ ]内时要换成$\left(\nabla-(\nabla\bar{r})\dfrac{\partial}{c\partial t}\right)$. 现在来求$\nabla^2 v=\nabla\cdot\nabla v$,

$$\begin{aligned}\nabla^2 v&=\nabla\cdot\left[\left(\nabla-(\nabla\bar{r})\frac{\partial}{c\partial t}\right)u\right]\\&=\nabla\cdot\left[\left(\nabla+(\nabla\bar{r})\frac{\partial}{c\partial t}\right)u-2(\nabla\bar{r})\frac{\partial u}{c\partial t}\right]\\&=\left[\left(\nabla-(\nabla\bar{r})\frac{\partial}{c\partial t}\right)\cdot\left(\nabla+(\nabla\bar{r})\frac{\partial}{c\partial t}\right)u\right]-2\nabla\cdot\left[(\nabla\bar{r})\frac{\partial u}{c\partial t}\right].\end{aligned}$$

注意到$\nabla\bar{r}=\dfrac{\bar{\boldsymbol{r}}}{\bar{r}}$,$(\nabla\bar{r})\cdot(\nabla\bar{r})=1$,以及

$$\nabla\cdot\left((\nabla\bar{r})\frac{\partial u}{c\partial t}\right)-\left((\nabla\bar{r})\frac{\partial}{c\partial t}\right)\cdot\nabla u=(\nabla^2\bar{r})\frac{\partial u}{c\partial t};$$

因而

$$\begin{aligned}\nabla^2 v&=\left[\left(\nabla^2-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)u\right]+(\nabla^2\bar{r})\left[\frac{\partial u}{c\partial t}\right]-2\nabla\cdot\left\{(\nabla\bar{r})\left[\frac{\partial u}{c\partial t}\right]\right\}\\&=[-f]-2\bar{r}\nabla\cdot\left[\frac{1}{\bar{r}}(\nabla\bar{r})\frac{\partial u}{c\partial t}\right];\end{aligned}$$

最后一步应用了波动方程,以及$-2\bar{r}\left(\nabla\dfrac{1}{\bar{r}}\right)\cdot\nabla\bar{r}=\dfrac{2}{\bar{r}}=\nabla^2\bar{r}$.

因为∇^2 是自伴算符,Green 定理给出

① H. Bateman, *Partial Differrential Equations of Mathematical Physics*, Cambridge, New York, 1944, p. 184.

$$\int_V \left\{ \frac{1}{\bar{r}} \nabla^2 v - v \nabla^2 \left(\frac{1}{\bar{r}} \right) \right\} \mathrm{d}\tau = \int_\Sigma \left\{ \frac{1}{\bar{r}} \nabla v - v \nabla \left(\frac{1}{\bar{r}} \right) \right\} \cdot \boldsymbol{n} \mathrm{d}\sigma.$$

利用上述 $v, \nabla v, \nabla^2 v$ 各式，并注意到 $\nabla^2 \left(\frac{1}{\bar{r}} \right) = 0$，上式可化为

$$\int_V \frac{[f]}{\bar{r}} \mathrm{d}\tau - \int_\Sigma \left\{ [u] \nabla \left(\frac{1}{\bar{r}} \right) - \frac{1}{\bar{r}} [\nabla u] - \left[\frac{\partial u}{c \partial t} \right] \frac{1}{\bar{r}} \nabla \bar{r} \right\} \cdot \boldsymbol{n} \mathrm{d}\sigma = 0.$$

应注意，所取积分区域 V 要将 $\boldsymbol{r}_1$ 点挖去. 为此将 Σ 变为 $\Sigma + \Sigma_1$，Σ_1 只包括 $\boldsymbol{r}_1$ 在内，取为一小球，$\boldsymbol{n}_1 (= -\boldsymbol{n})$ 由 $\boldsymbol{r}_1$ 指向 V. 对于 Σ_1 面上的积分，由于 $\mathrm{d}\sigma \to 4\pi \bar{r}^2$，$\nabla \left(\frac{1}{\bar{r}} \right) \cdot \boldsymbol{n} = \frac{1}{\bar{r}^2}$；当 $\bar{r} \to 0$ 时，{} 中第一项的积分为 $4\pi u(\boldsymbol{r}_1, t)$，其余积分为零. 所以，上式的最后结果可写成

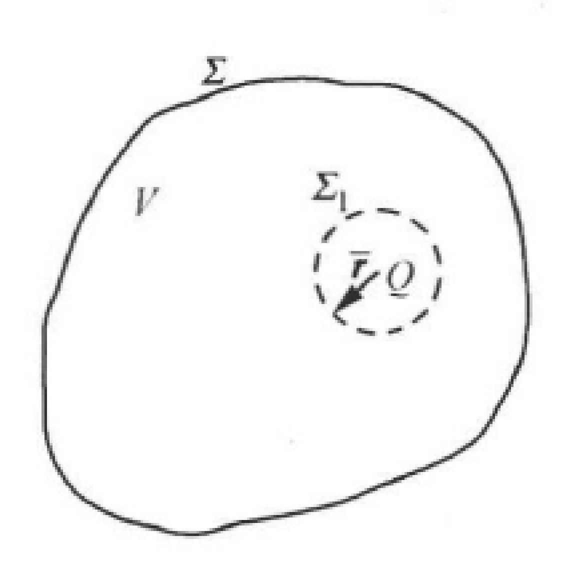
$$\begin{aligned} u(\boldsymbol{r}_1, t) = & \frac{1}{4\pi} \int_V \frac{[f(\boldsymbol{r}, t)]}{|\boldsymbol{r} - \boldsymbol{r}_1|} \mathrm{d}\tau \\ & - \frac{1}{4\pi} \int_\Sigma \left\{ [u(\boldsymbol{r}, t)] \nabla \frac{1}{|\boldsymbol{r} - \boldsymbol{r}_1|} - \frac{1}{|\boldsymbol{r} - \boldsymbol{r}_1|} [\nabla u(\boldsymbol{r}, t)] \right. \\ & \left. - \left[\frac{\partial u(\boldsymbol{r}, t)}{c \partial t} \right] \frac{1}{|\boldsymbol{r} - \boldsymbol{r}_1|} \nabla |\boldsymbol{r} - \boldsymbol{r}_1| \right\} \cdot \boldsymbol{n} \mathrm{d}\sigma. \end{aligned}$$

这就是所要证明的 Kirchhoff 公式.

这是一个定理，它将 D'Alembert 方程

$$\square^2 u = -f$$

的解用推迟的体积分和面积分表达. 如下图所示，区域 V 内任一点 $\boldsymbol{r}_1$ 之势 $u(\boldsymbol{r}_1, t)$，决定于 V 内诸点 $\boldsymbol{r}$ 处体源推迟值 $[f(\boldsymbol{r}, t)]$ 以及外边界 Σ 上诸点 $\boldsymbol{r}$ 处之推迟值 $[u(\boldsymbol{r}, t)]$，$\left[\frac{\partial}{\partial n} u(\boldsymbol{r}, t) \right]$，$\left[\frac{1}{c} \frac{\partial u}{\partial t} \right]$，即[ ]内各量在 $t - \frac{\bar{r}}{c}$ 时刻之值，其中 $\bar{r} = |\boldsymbol{r} - \boldsymbol{r}_1|$.



就数学而言，也可将方程的解用超前的体积分和面积分表达，即用 $t + \frac{\bar{r}}{c}$ 时刻各量之值来决定 $u(\boldsymbol{r}_1, t)$. 然而，就物理上的边界条件而言，则有不同看法：只有过去决定将来而不是将来决定现在. 因此，只有推迟解才是物理上合理的.

11.3.4 波动方程的叠加原理解法

下面来讨论波动方程

$$Lu=\left(\nabla^2-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)u=0$$

的具体解法. 首先用分离变量法求特解，然后用这些特解来构造通解；最后再由边界条件来确定通解中的常数，以求得适合实际问题之解. 这里将只讨论其中前一部分.

(1) 分离变量法求平面波解

由于波动方程

$$\Box^2 u=\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)u=0$$

中微分算符 L 的几部分互不相关，所以可用**分离变量法**进行求解.

先假定一个特解 u_1 为

$$u_1=S(x,y,z)T(t),$$

代入原方程后得

$$\nabla^2 u_1=(\nabla^2 S)T=\frac{1}{c^2}\frac{\partial^2 u_1}{\partial t^2}=S\frac{1}{c^2}\frac{\mathrm{d}^2 T}{\mathrm{d}t^2},$$

或者写成

$$\underbrace{\frac{\nabla^2 S}{S}=\frac{\frac{1}{c^2}\frac{\mathrm{d}^2 T}{\mathrm{d}t^2}}{T}}_{=-k^2},$$

因为按假定，左边仅为 x,y,z 的函数，右边仅为 t 的函数，而 x,y,z,t 均为自变量，不可能存在函数关系；故要求它们都等于常数 $-k^2$，后者称为**分离常数**. 于是有

$$\begin{cases}\nabla^2 S+k^2 S=0,\\ \dfrac{1}{c^2}\dfrac{\mathrm{d}^2 T}{\mathrm{d}t^2}+k^2 T=0.\end{cases}$$

如果以上方程有解，则所作假定是对的，否则不对.

因为 Laplace 符 ∇^2 也是由互不相关的几个部分所组成，再次应用分离变量法，令

$$S(x,y,z)=X(x)Y(y)Z(z),$$

于是上面第一式可变为

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)XYZ+k^2 XYZ=0,$$

$$YZ\frac{\mathrm{d}^2X}{\mathrm{d}x^2}+ZX\frac{\mathrm{d}^2Y}{\mathrm{d}y^2}+XY\frac{\mathrm{d}^2Z}{\mathrm{d}z^2}+k^2XYZ=0;$$

或者写成

$$\underset{\underset{-k_x^2}{\|}}{\frac{\frac{\mathrm{d}^2X}{\mathrm{d}x^2}}{X}}+\underset{\underset{-k_y^2}{\|}}{\frac{\frac{\mathrm{d}^2Y}{\mathrm{d}y^2}}{Y}}+\underset{\underset{-k_z^2}{\|}}{\frac{\frac{\mathrm{d}^2Z}{\mathrm{d}z^2}}{Z}}+k^2=0;$$

同上理由，$-k_x^2$，$-k_y^2$，$-k_z^2$ 亦称分离常数，并且有

$$k_x^2+k_y^2+k_z^2=k^2.$$

因为否则的话，如果它们分别等于 $f(x)$，$g(y)$，$h(z)$，要求对一切可能的 x,y,z 都满足

$$f(x)+g(y)+h(z)+k^2=0$$

是不可能的.

所以，最后得到

$$\begin{cases}\dfrac{\mathrm{d}^2X}{\mathrm{d}x^2}+k_x^2X=0,\\[2mm]\dfrac{\mathrm{d}^2Y}{\mathrm{d}y^2}+k_y^2Y=0,\\[2mm]\dfrac{\mathrm{d}^2Z}{\mathrm{d}z^2}+k_z^2Z=0,\\[2mm]\dfrac{1}{c^2}\dfrac{\mathrm{d}^2T}{\mathrm{d}t^2}+k^2T=0,\\[2mm]k_x^2+k_y^2+k_z^2=k^2.\end{cases}$$

每个方程的解都具有类似形式，即

$$\begin{cases}X=\mathrm{e}^{\mathrm{i}k_xx}\quad\text{或}\quad\mathrm{e}^{-\mathrm{i}k_xx},\\Y=\mathrm{e}^{\mathrm{i}k_yy}\quad\text{或}\quad\mathrm{e}^{-\mathrm{i}k_yy},\\Z=\mathrm{e}^{\mathrm{i}k_zz}\quad\text{或}\quad\mathrm{e}^{-\mathrm{i}k_zz},\\T=\mathrm{e}^{\mathrm{i}\omega t}\quad\text{或}\quad\mathrm{e}^{-\mathrm{i}\omega t};\end{cases}$$

其中 $\omega=kc$. 当然，也可采用例如 $X=\cos k_xx$ 或 $\sin k_xx$ 等的形式.

于是，波动方程的一个特解是

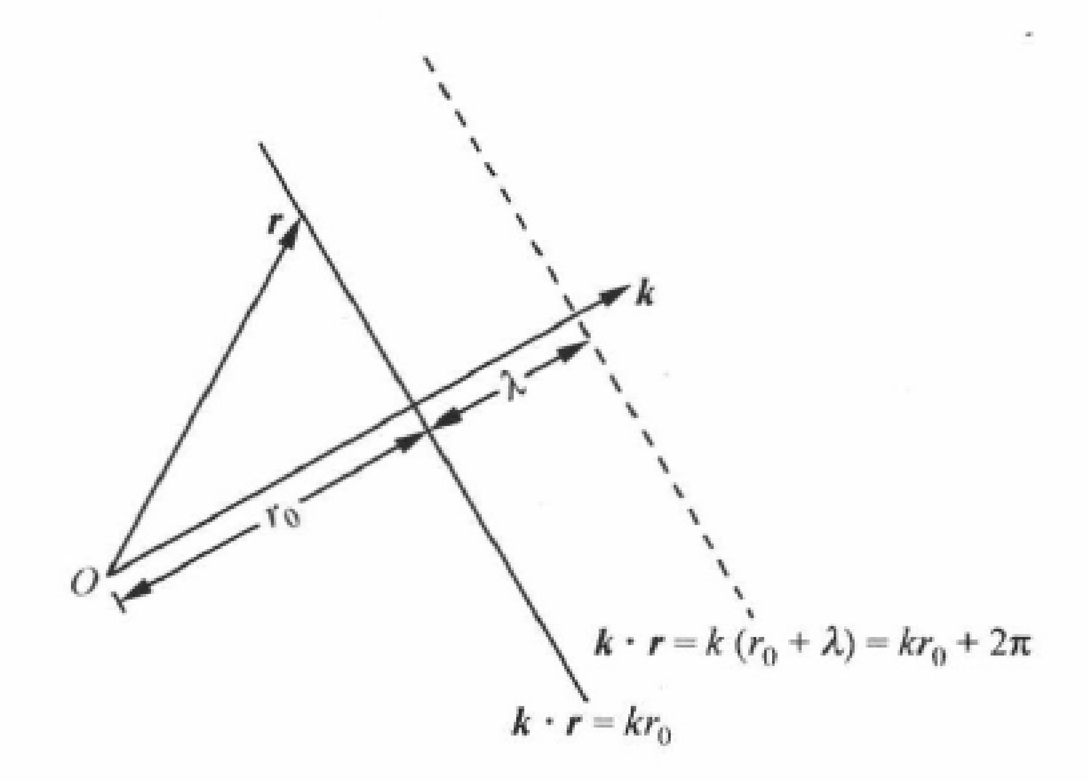
$$u_1=\mathrm{e}^{\mathrm{i}(k_xx+k_yy+k_zz-\omega t)}=\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)},$$

其中 $\boldsymbol{k}=(k_x,k_y,k_z)$，$\boldsymbol{r}=(x,y,z)$.

由解析几何知道(如下页图)，

$$\boldsymbol{k}\cdot\boldsymbol{r}=k_xx+k_yy+k_zz=\text{const}$$

是空间中平面的方程.



由此，在某一时刻 t，$\boldsymbol{k}\cdot\boldsymbol{r}=kr_0$（$=\text{const}$）的平面上，处处相位相同，即处处 u_1 相同. 同时，在 $\boldsymbol{k}\cdot\boldsymbol{r}=k(r_0+\lambda)=kr_0+2\pi$ 的另一平面上，有 $u_1(r_0+\lambda)=u_1(r_0)$；即相隔为 λ 的二平面是同相的. 所以，u_1 表示**平面波**，$\boldsymbol{k}\cdot\boldsymbol{r}=\text{const}$ 的平面称为**波阵面**，而 $\lambda=\dfrac{2\pi}{k}$则称为**波长**，而 $k=\dfrac{2\pi}{\lambda}$则称为**波数**. $\boldsymbol{k}$ 称为**波矢**，它垂直于波阵面.

再看 u_1 随 t 的变化. 就一固定波阵面而言，u_1 以**周期**为$\dfrac{2\pi}{\omega}$而随 t 发生变化；ω 称为**角频率**. 另一方面，若保持 $\boldsymbol{k}\cdot\boldsymbol{r}-\omega t=k(r_0-ct)=\text{const}$，则 u_1 不变，即波阵面以 $r_0=ct+\text{const}$ 而行进，c 是**波速**，$\boldsymbol{k}$ 是波行进方向. 所以，这样的特解 u_1 代表**平面行波**.

顺便提一下，二行波可叠加而成一**驻波**. 驻波的意思是，**波节**处永远是波节，**波腹**处永远是波腹，停驻不前；如$\sin k_x x\ \sin k_y y\ \sin k_z z\ \sin\omega t$ 即是一例.

每一个特解 u_1 代表一个平面行波，它由三个常数 k_x,k_y,k_z 为表征. 由于三个常数可以自由变，所以总共有∞^3 个特解.

凡线性微分方程，若知其全部特解，则可采用叠加原理求其通解. 因此，对于波动方程

$$\nabla^2 u-\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}=0,$$

其通解是由∞^3 个平面行波解叠加而成，即

$$\begin{aligned}u(\boldsymbol{r},t)&=\iiint A(k_x,k_y,k_z)\exp[\mathrm{i}(k_x x+k_y y+k_z z-\omega t)]\mathrm{d}k_x\mathrm{d}k_y\mathrm{d}k_z\\&=\int A(\boldsymbol{k})\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\mathrm{d}\boldsymbol{k}.\end{aligned}$$

(2) 球面坐标系中的波动方程

下面在球面坐标系

$$\begin{cases}x=r\sin\theta\cos\varphi,\\y=r\sin\theta\sin\varphi,\\z=r\cos\theta\end{cases}$$

中来求解波动方程

$$\nabla^2 u - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0.$$

仍采用分离变量法来做. 首先,令

$$u_1 = S\mathrm{e}^{\pm \mathrm{i}\omega t},$$

得到($\omega = kc$)

$$\nabla^2 S + k^2 S = 0.$$

采用球面坐标系中 Laplace 符∇^2 的表达式(见第 9 章 9.6.4 节之(1)),得到

$$\left\{\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] + k^2\right\}S = 0,$$

这里 $S = S(r,\theta,\varphi)$. 再令

$$S(r,\theta,\varphi) = R(r)\mathrm{Y}(\theta,\varphi),$$

得到

$$\frac{r^2\left[\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + k^2\right]R(r)}{R(r)} = l(l+1),$$

$$\frac{\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]\mathrm{Y}(\theta,\varphi)}{\mathrm{Y}(\theta,\varphi)} = -l(l+1);$$

其中 $l(l+1)$是分离常数. 最后令

$$\mathrm{Y}(\theta,\varphi) = \Theta(\theta)\Phi(\varphi),$$

又可将后一方程分离成二方程.

所以,在球面坐标下,波动方程变为

$$\begin{cases} \left\{\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}}{\mathrm{d}r} - \frac{l(l+1)}{r^2} + k^2\right\}R = 0, \\ \left\{\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}}{\mathrm{d}\theta}\right) - \frac{m^2}{\sin^2\theta} + l(l+1)\right\}\Theta = 0, \\ \left(\frac{\mathrm{d}^2}{\mathrm{d}\varphi^2} + m^2\right)\Phi = 0, \\ \left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + k^2c^2\right)T = 0; \end{cases}$$

有三个分离常数 $k^2, l(l+1), m^2$.

后两个方程极易求解,不必赘述. 径向方程可化为**球 Bessel 方程**[令 $\zeta = kr$, $\psi(\zeta) = R(r)$即得,参见第 12 章 12.5.5 节];而关于 θ 的方程可化为**连带 Legendre 方程**[令 $\mu = \cos\theta, v(\mu) = \Theta(\theta)$即得,参见第 12 章 12.4.3 节之(1)和第 12 章 12.4.4 节].

(3) 柱面坐标系中的波动方程

柱面坐标系(ρ,φ,z)中的波动方程是[参见第 9 章 9.6.4 节之(2)]

$$\frac{\partial^2 u}{\partial \rho^2}+\frac{1}{\rho}\frac{\partial u}{\partial \rho}+\frac{1}{\rho^2}\frac{\partial^2 u}{\partial \varphi^2}+\frac{\partial^2 u}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}=0.$$

采用分离变量法，令 $u(\rho,\varphi,z,t)=R(\rho)\Phi(\varphi)Z(z)T(t)$，类似地容易得到

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}\rho^2}+\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}+\left(k_\perp^2-\frac{m^2}{\rho^2}\right)\right\}R=0,$$

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}\varphi^2}+m^2\right\}\Phi=0,$$

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}z^2}+k_z^2\right\}Z=0,$$

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}t^2}+k^2c^2\right\}T=0;$$

有三个分离常数 k^2,k_z^2,m^2，而 $k_\perp^2=k^2-k_z^2$.

后三个方程极易求解，不必赘述. 径向方程可化为 **Bessel 方程**［令 $\zeta=k_\perp\rho,\psi(\zeta)=R(\rho)$ 即得，参见第 12 章 12.5.5 节］.

由以上作法可知，用分离变量法可将一个偏微分方程变成几个常微分方程.

第 12 章　二阶线性常微分方程

12.1 引　论

本章将专门论述形式为

$$\frac{\mathrm{d}^2 u}{\mathrm{d}z^2}+p(z)\frac{\mathrm{d}u}{\mathrm{d}z}+q(z)=0$$

的二阶线性常微分方程的解.

首先,介绍一下关于解的一般知识.

12.1.1 解的基本概念

先用几个简单例子来说明有关解的若干基本概念.

考虑下列线性微分方程

$$\frac{\mathrm{d}^2 u}{\mathrm{d}z^2}+m^2 u=0.$$

显然,$u_1=\cos mz$ 和 $u_2=\sin mz$ 都满足此方程,称为方程的**特解**.同时,注意到 u_1 和 u_2 是相互独立的,此二特解的线性组合

$$u=c_1u_1+c_2u_2=c_1\cos mz+c_2\sin mz,$$

其中 c_1 和 c_2 是任意常数,也是方程的解;称为方程的**通解**.另一方面,$u_+=\mathrm{e}^{\mathrm{i}mz}$ 和 $u_-=\mathrm{e}^{-\mathrm{i}mz}$ 也是方程相互独立之特解,$u=c_+\mathrm{e}^{\mathrm{i}mz}+c_-\mathrm{e}^{-\mathrm{i}mz}$ 同样构成方程的通解;而且与前面的通解是等价的.因此,任何相互独立的两个特解 u_1,u_2 称为方程的**基本解组**.

对于非线性微分方程,例如,

$$\left(\frac{\mathrm{d}^2 u}{\mathrm{d}z^2}\right)^2-z\frac{\mathrm{d}^2 u}{\mathrm{d}z^2}+\frac{\mathrm{d}u}{\mathrm{d}z}=0,$$

可求得两个特解为 $u_1=z^2-4z$ 和 $u_2=1$,而其通解则为

$$u=c_1z^2-4c_1^2z+c_2,$$

仍含两个任意常数,但并非上述特解之线性组合.另外,还有一个解 $u_s=\frac{1}{12}z^3$ 不包括在上述通解之内,称为方程之**奇解**.

对于一般常微分方程的解,有下述定理(可参考 10.3.3 节以及 12.1.3 节):

定理　一个 n 阶常微分方程的通解(解的总体,可能的奇解除外)包含 n 个任意积分常数 $c_1,c_2,\cdots,c_n$,即 $u=u(z;c_1,\cdots,c_n)$;若微分方程是线性的,则其通解对

$c_1, c_2, \cdots, c_n$ 也是线性的，即 $u = \sum_{i=1}^{n} c_i u_i(z)$，并且无奇解，而 n 个独立特解 $u_1, u_2, \cdots, u_n$ 构成微分方程的基本解组.

12.1.2　降阶法

对于二阶线性常微分方程

$$\frac{\mathrm{d}^2 u}{\mathrm{d}z^2} + p(z)\frac{\mathrm{d}u}{\mathrm{d}z} + q(z)u = 0,$$

如果已经知道一个特解 $u_1(z)$，则可令

$$u(z) = u_1(z)v(z) \quad 和 \quad \frac{\mathrm{d}v(z)}{\mathrm{d}z} = w(z),$$

代入原方程后可得

$$\begin{aligned} 0 &= u_1\left(\frac{\mathrm{d}^2 v}{\mathrm{d}z^2} + p(z)\frac{\mathrm{d}v}{\mathrm{d}z}\right) + 2\left(\frac{\mathrm{d}u_1}{\mathrm{d}z}\right)\left(\frac{\mathrm{d}v}{\mathrm{d}z}\right) + \left(\frac{\mathrm{d}^2 u_1}{\mathrm{d}z^2} + p(z)\frac{\mathrm{d}u_1}{\mathrm{d}z} + q(z)u_1\right)v \\ &= u_1\left(\frac{\mathrm{d}w}{\mathrm{d}z} + \left[p(z) + \frac{2}{u_1}\frac{\mathrm{d}u_1}{\mathrm{d}z}\right]w\right). \end{aligned}$$

因为 $u_1 \neq 0$，结果可将二阶微分方程降阶为下列一阶微分方程

$$\frac{\mathrm{d}w}{\mathrm{d}z} + P(z)w = 0,$$

其中 $P(z) = p(z) + \frac{2}{u_1}\frac{\mathrm{d}u_1}{\mathrm{d}z}$；此法称为**降阶法**.

此一阶微分方程很容易求解，因为

$$\frac{\mathrm{d}w}{w} + P(z)\mathrm{d}z = 0 \to w = w_0 \exp\left[-\int_{z_0}^{z} P\mathrm{d}z\right],$$

其中 $w_0 = w(z_0)$. 又由 $\frac{\mathrm{d}v}{\mathrm{d}z} = w$ 积分可得

$$v = \int_{z_0}^{z} w\mathrm{d}z + v_0, \quad v_0 = v(z_0).$$

最后求得二阶微分方程的解为

$$u = u_1\left\{v_0 + w_0\int_{z_0}^{z} \exp\left[-\int_{z_0}^{\zeta} P(\xi)\mathrm{d}\xi\right]\mathrm{d}\zeta\right\}.$$

若令 $c_1 = v_0, c_2 = w_0, u_2 = u_1\int_{z_0}^{z} \exp\left[-\int_{z_0}^{\zeta} P(\xi)\mathrm{d}\xi\right]\mathrm{d}\zeta$，则

$$u = c_1 u_1 + c_2 u_2.$$

也就是说，对于二阶线性常微分方程，如果知道它的一个特解 u_1，则另一个特解 u_2 就可以用求积分的方法将其求出，同时也可知微分方程的通解具有如上的形式.

上述降阶法很容易推广到一般的 n 阶线性常微分方程，可表述为如下定理.

定理　若 $u_1(z)$是 n 阶线性齐次常微分方程

$$\frac{\mathrm{d}^n u}{\mathrm{d}z^n} + p_1(z)\frac{\mathrm{d}^{n-1} u}{\mathrm{d}z^{n-1}} + \cdots + p_n(z)u = 0$$

的一个特解,则通过令

$$u(z) = u_1(z)v(z) \quad \text{和} \quad \frac{\mathrm{d}v}{\mathrm{d}z} = w(z),$$

可将 n 阶线性非齐次常微分方程

$$\frac{\mathrm{d}^n u}{\mathrm{d}z^n} + p_1(z)\frac{\mathrm{d}^{n-1} u}{\mathrm{d}z^{n-1}} + \cdots + p_n(z)u = r(z),$$

变换为$(n-1)$阶线性常微分方程

$$\frac{\mathrm{d}^{n-1} w}{\mathrm{d}z^{n-1}} + P_1(z)\frac{\mathrm{d}^{n-2} w}{\mathrm{d}z^{n-2}} + \cdots + P_{n-1}(z)w = r(z).$$

12.1.3　初值问题的另一看法

对于常微分方程

$$\frac{\mathrm{d}^2 u}{\mathrm{d}z^2} + p(z)\frac{\mathrm{d}u}{\mathrm{d}z} + q(z)u = 0,$$

在 $z=z_0$ 时给定

$$u(z_0) = u_0, \quad \frac{\mathrm{d}u}{\mathrm{d}z}(z_0) = u_0',$$

求在 z_0 点邻域的解 $u(z)$;这样一个**初值问题**,还可有另一种看法.

首先,将方程改写为

$$\frac{\mathrm{d}^2 u}{\mathrm{d}z^2} = - p(z)\frac{\mathrm{d}u}{\mathrm{d}z} - q(z)u = f\left(z, u, \frac{\mathrm{d}u}{\mathrm{d}z}\right),$$

上式右边在 $z=z_0$ 点之值均为已知,由此可求得$\frac{\mathrm{d}^2 u}{\mathrm{d}z^2}$在 $z=z_0$ 之值.

将上式对 z 求导一次,记作

$$\frac{\mathrm{d}^3 u}{\mathrm{d}z^3} = f_1\left(z, u, \frac{\mathrm{d}u}{\mathrm{d}z}, \frac{\mathrm{d}^2 u}{\mathrm{d}z^2}\right),$$

注意到右边所含高阶导数总比左边低一阶. 现在,上式右边在 $z=z_0$ 点之值亦均为已知,由此可求得$\frac{\mathrm{d}^3 u}{\mathrm{d}z^3}$在 $z=z_0$ 之值.

这样继续进行下去,就可得各阶导数在 $z=z_0$ 点之值. 也就是说,**初值问题的微分方程可当做计算高阶导数的方程**.

于是,可求得微分方程在 z_0 点邻域的 Taylor 级数解为

$$u(z) = u(z_0) + (z - z_0)\left(\frac{\mathrm{d}u}{\mathrm{d}z}\right)_{z_0} + \frac{1}{2}(z - z_0)^2\left(\frac{\mathrm{d}^2 u}{\mathrm{d}z^2}\right)_{z_0} + \cdots.$$

显然可见，以上作法同样适用于一般的 n 阶常微分方程

$$\frac{\mathrm{d}^n u}{\mathrm{d}z^n} = f\left(z, u, \frac{\mathrm{d}u}{\mathrm{d}z}, \cdots, \frac{\mathrm{d}^{n-1} u}{\mathrm{d}z^{n-1}}\right)$$

的初值问题.

这就是微分方程初值问题的另一种意义.

应该指出，上面的论述未曾考虑 $u(z)$ 能否不断地求导下去的问题. 如果 m 阶导数 $\frac{\mathrm{d}^m u}{\mathrm{d}z^m}$ 不能再求导，就不能求得（$>m$ 的）更高阶导数之值，因而 $u(z)$ 在 z_0 邻近的值也就不得而知了.

人们知道，复变函数中的解析函数可以无限地求导下去. 所以，若 u 是解析函数，则上面的做法是不成问题的. 因此，在讨论微分方程时，总要假定 u 是解析函数. 这样就需要知道一些有关复变函数性质的知识，将在下一节概要地予以介绍.

12.1.4　函数的级数表示和积分表示

一般而言，函数主要有两种表示方法：

$$\begin{cases}\text{级数表示,}\\ \text{积分表示.}\end{cases}$$

具体来说，例如：

$$\begin{cases}\text{幂级数表示（}z^n\text{ 的级数，}n\text{ 取整数），}\\ \text{Fourier 级数或 Fourier 积分表示}\\ \quad\text{（}\mathrm{e}^{\mathrm{i}nz}\text{ 的级数，前者 }n\text{ 取整数，后者 }n\text{ 取连续值）.}\end{cases}$$

它们的特性是：

$$\begin{cases}\text{幂级数可以逐项微分，用在复变函数；}\\ \text{Fourier 级数可以逐项积分，用在实变函数.}\end{cases}$$

所以，微分方程中用幂级数较多.

12.2　复变函数论概要

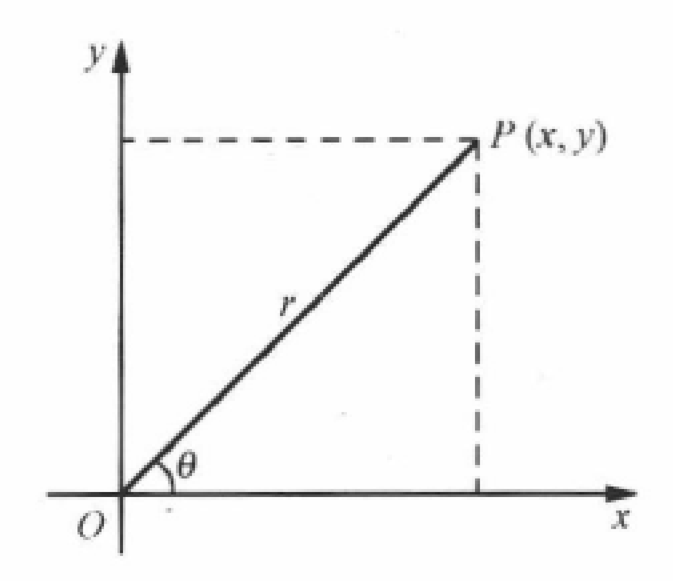
12.2.1　解析函数的定义

(1) 复数

复数 $z=x+\mathrm{i}y$，其中 x, y 是实数，$\mathrm{i}=\sqrt{-1}$ 是**虚数单位**.

实数可用直线上的点表示；复数则可用平面上的点表示如下页图，称之为 **Argand 图**. 由图显然可见，

$$z = x + \mathrm{i}y = r(\cos\theta + \mathrm{i}\sin\theta) = r\mathrm{e}^{\mathrm{i}\theta}.$$



复数 z 的**实部**和**虚部**分别记作

$$\operatorname{Re} z = x = r\cos\theta, \quad \operatorname{Im} z = y = r\sin\theta.$$

复数 z 的**模**和**辐角**则分别记作

$$|z| = r = \sqrt{x^2 + y^2}, \quad \arg z = \theta = \arctan\frac{y}{x}.$$

与复数 z 为复共轭的数称为其**共轭复数**,记作

$$z^* = x - \mathrm{i}y.$$

显然可见,$z^* z = |z|^2$.

应该指出,复数的相等当且仅当其实部与虚部分别相等,即

$$z_1 = z_2 \iff \operatorname{Re}z_1 = \operatorname{Re}z_2,\ \operatorname{Im}z_1 = \operatorname{Im}z_2.$$

另一方面,复数的加法和乘法显然为

$$z_1 + z_2 = (x_1 + \mathrm{i}y_1) + (x_2 + \mathrm{i}y_2) = (x_1 + x_2) + \mathrm{i}(y_1 + y_2),$$

$$z_1 \cdot z_2 = (x_1 + \mathrm{i}y_1)(x_2 + \mathrm{i}y_2) = (x_1x_2 - y_1y_2) + \mathrm{i}(x_1y_2 + x_2y_1).$$

前面(第 1 章 1.1A 节)曾经指出,全体复数构成复数域 **C**;这一点很容易证明,此处从略.

顺便指出,复数域是最复杂的数域①.

(2) 复变函数

以复数为自变量(即**复变数**)的复值函数称为**复变函数**,可表示为

---

① 将复数再进一步推广就是以 $1, i, j, k$ 为**基**的**四元数**,

$$q = t + xi + yj + zk,$$

其中 t, x, y, z 是实数,基单位 i, j, k 满足下列关系

$$i^2 = j^2 = k^2 = -1,$$

$$ij = k = -ji,\ jk = i = -kj,\ ki = j = -ik.$$

但是,四元数全体不能构成域,因为它不满足乘法的交换律,也不满足可除律(若 $ab=0$,则有或者 $a=0$,或者 $b=0$).然而,它可以构成环.

环是定义了加法和乘法这两种二元运算的集合 R,关于加法是一个交换群,且乘法对加法满足分配律.

因为对乘法未作限制,所以,环是比域更为广义的代数结构.

如果乘法满足结合律,则称为**结合环**或**环**,否则为**非结合环**.若环 R 中含有单位元素,且每个非零元素都有逆元素,则称为**除环**.四元数全体构成典型的除环.若在除环的定义中再规定其乘法还满足交换律,则为域,即域为**交换除环**.

可参考:《中国大百科全书:数学》,304 页,环(谢邦杰撰),中国大百科全书出版社,北京,1988.

$$w(z) = w(x+\mathrm{i}y) = u(x,y) + \mathrm{i}v(x,y),$$

即一个复变函数 $w(z)$ 对应着一对两个实变量的实值函数. 它同样可以写成

$$w = R\mathrm{e}^{\mathrm{i}\Theta},$$

其中 $R=\sqrt{u^2+v^2}$ 和 $\Theta=\arctan\left(\frac{v}{u}\right)$ 分别为 $w(z)$ 的模和辐角. 例如,

$$w = z^2 = x^2 - y^2 + 2\mathrm{i}xy,$$

$$R = |w| = x^2 + y^2, \quad \Theta = \arctan\frac{2xy}{x^2-y^2}.$$

(3) 解析函数

复变函数论中研究的主要对象是解析函数,它们是一类具有特殊性质的复变函数.

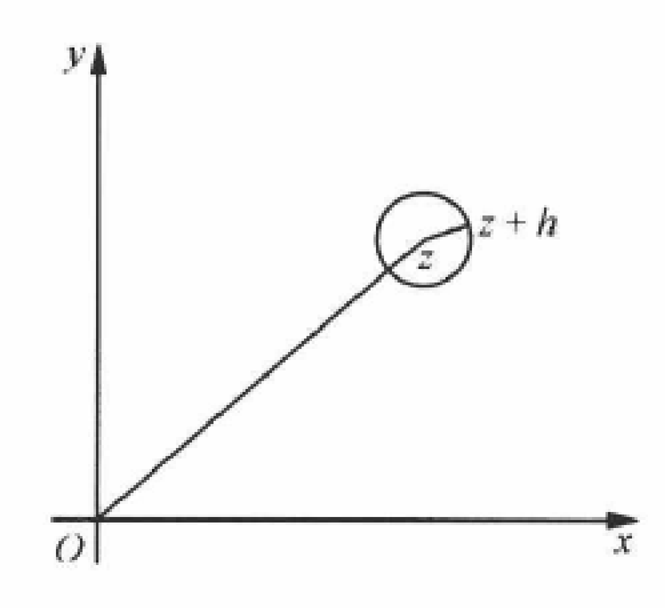
下面首先给出 **Cauchy 关于解析函数的定义.**

对于一个复变函数 $f(z)$,若

$$\lim_{h\to 0}\frac{f(z+h)-f(z)}{h}\ \text{存在,且与}\ \arg h\ \text{无关},$$

则 $f(z)$ 是 z 的**解析函数.**

所谓极限与 $\arg h$ 无关指的是,如下图,无论从圆上哪一点趋近 z,极限总是一样. 此极限即定义为复变函数 $f(z)$ 的导数,记作$\frac{\mathrm{d}f}{\mathrm{d}z}$.



例 1 $f(z)=\mathrm{e}^z$ 是解析函数,因为

$$\frac{\mathrm{e}^{z+h}-\mathrm{e}^z}{h} = \frac{\mathrm{e}^z}{h}(\mathrm{e}^h - 1) = \mathrm{e}^z\left(1+\frac{h}{2!}+\cdots\right),$$

$$\lim_{h\to 0}\frac{\mathrm{e}^{z+h}-\mathrm{e}^z}{h} = \mathrm{e}^z, \quad \text{即} \quad \frac{\mathrm{d}f}{\mathrm{d}z} = \mathrm{e}^z.$$

例 2 $f(z)=x-\mathrm{i}y=z^*$ 不是解析函数,因为若令 $h=k+\mathrm{i}l$,有 $\arg h=\arctan\frac{l}{k}$ 以及

$$\frac{f(z+h)-f(z)}{h} = \frac{(z^*+h^*)-z^*}{h} = \frac{h^*}{h} = \frac{k-\mathrm{i}l}{k+\mathrm{i}l}$$

$$=\begin{cases}1, & 若\ \arg h=0,\\ -1, & 若\ \arg h=\dfrac{\pi}{2};\end{cases}$$

因而其极限依赖于 $\arg h$ 而有所不同，故 $f(z)=z^*$ 非解析函数.

同样可证明 $f(z)=|z|,|z|^2(=zz^*)$都不是解析函数.

例 3 $f(z)=\dfrac{1}{z-a}$,

$$\frac{1}{h}[f(z+h)-f(z)]=\frac{1}{h}\left[\frac{1}{z+h-a}-\frac{1}{z-a}\right]=-\frac{1}{(z-a)(z+h-a)};$$

当 $z\neq a$ 时有 $\dfrac{\mathrm{d}f}{\mathrm{d}z}=-\dfrac{1}{(z-a)^2}$，而当 $z=a$ 时则 $\dfrac{\mathrm{d}f}{\mathrm{d}z}$ 不存在. 因此，$f(z)=\dfrac{1}{z-a}$ 对除 $z=a$ 这一点之外的其他 z 值而言是解析函数，而 $z=a$ 这点则称为函数 $f(z)$ 的**奇点**.

12.2.2 函数的奇点与支点

(1) 函数的奇点

下面进一步来讨论使函数不可求导的奇点. 先假定 $f(z)$ 是单值函数，即假定对每个 z 值有唯一的 $f(z)$ 值相对应.

单值函数 $f(z)$ 的**奇点**可分为两类：

(i) **极点** 例如，$\dfrac{1}{(z-a)^n}$ 具有 n 阶极点. 一般而言，凡可展成下列形式的函数，

$$f(z)=\frac{A_{-n}}{(z-a)^n}+\frac{A_{-(n-1)}}{(z-a)^{n-1}}+\cdots+\frac{A_{-1}}{(z-a)}+A_0+A_1(z-a)+\cdots,$$

$A_{-n}\neq 0$ 时，$z=a$ 点称为函数的 n **阶极点**. 前面曾经提过的 $f(z)=\dfrac{1}{z-a}$，$z=a$ 点是其**一阶极点**，常称为**单极点**.

(ii) **本质奇点** 除极点以外的任何奇点都是本质奇点. 特别是，可展成含 $(z-a)^{-1}$ 之无穷级数的函数，

$$f(z)=\sum_{n=-\infty}^{m}A_n(z-a)^n,$$

$z=a$ 点为函数的本质奇点. 例如，

$$f(z)=\exp\left(\frac{1}{z-a}\right)=\sum_{n=0}^{\infty}\frac{1}{n!}(z-a)^{-n}.$$

应该指出，以上举例是指**孤立奇点**而言，因为 $f(z)$ 在除此点之外的充分小邻域内处处可导. 然而，还有另一类函数，例如，$\left(\sin\dfrac{1}{z}\right)^{-1}$，$z=\dfrac{1}{n\pi}$（$n$ 为整数）是其奇点；显然，极限点 $z=0$ 是函数的**非孤立奇点**. 所有非孤立奇点均为本质奇点.

现在可将单值函数的奇点分类如下：

$$\text{奇点}\begin{cases}\left.\begin{array}{l}\text{极点}\begin{cases}\text{单极点}\\ n\text{ 阶极点}\end{cases}\\ \text{本质奇点}\begin{cases}\text{孤立的(函数展式含有}(z-a)^{-1}\\ \qquad\text{的无穷级数)}\end{cases}\end{array}\right\}\text{孤立奇点}\\ \qquad\qquad\text{非孤立的}\qquad\qquad\qquad\qquad\}\text{非孤立奇点}\end{cases}$$

另外，关于函数 $f(z)$ 在无穷远点（$|z|\to\infty$）解析与否的问题，可通过变换

$$z_1=\frac{1}{z},\quad f(z)=F(z_1),$$

而视 $F(z_1)$ 在 $z=0$ 点处是否解析而定.

例 1　$f(z)=z^2\to F(z_1)=\dfrac{1}{z_1^2}$，$z_1=0$（即 $z=\infty$）为 2 阶极点. 一般而言，n 次多项式在无穷远点为 n 阶极点.

例 2　$f(z)=\mathrm{e}^z\to F(z_1)=\mathrm{e}^{\frac{1}{z_1}}=\displaystyle\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{1}{z_1}\right)^n$，所以，$z=\infty$ 点是 e^z 函数的本质奇点.

例 3　$f(z)=\dfrac{1}{\sin z}$，$z=n\pi$（$n=0,\pm1,\pm2,\cdots$）均为其单极点. 一般而言，若 $f(z)=\dfrac{1}{g(z)}$，则 $g(z)$ 的 m 重零点即为 $f(z)$ 的 m 阶极点.

例 4　$f(z)=\dfrac{\sin z}{z}$ 在 $z=0$ 点无定值，但若规定在此点取其级数展开之极限值，即 $f(0)=1$，则此函数可视为无奇性. 也有将这样的点称之为**可去奇点**.

(2) 多值函数的支点

对于复变数 $z=r\mathrm{e}^{\mathrm{i}\theta}$，由于 $\mathrm{e}^{2n\pi\mathrm{i}}=1$，所以，$z$ 的辐角具有不确定性；即可能为 θ，$2\pi+\theta$，$4\pi+\theta$，…，因此，在复变函数里引起了多值的问题.

例如，对于实变数：$\ln1=0$；而对于复变数：$\ln1=2n\pi\mathrm{i}$（$n=0,\pm1,\pm2,\cdots$）. 类似地，$\sqrt{1}=\sqrt{1\mathrm{e}^{2n\pi\mathrm{i}}}=|\sqrt{1}|\,\mathrm{e}^{n\pi\mathrm{i}}=+1$ 或 -1.

下面来看 $(z-a)^{\frac{1}{2}}$ 和 $\ln(z-a)$ 的情形.

$$z-a=|z-a|\,\mathrm{e}^{\mathrm{i}(\theta+2n\pi)},$$

$$(z-a)^{\frac{1}{2}}=|z-a|^{\frac{1}{2}}\mathrm{e}^{\mathrm{i}\left(\frac{\theta}{2}+n\pi\right)}=\pm|z-a|^{\frac{1}{2}}\mathrm{e}^{\mathrm{i}\frac{\theta}{2}},$$

$$\ln(z-a)=\ln|z-a|+\mathrm{i}\theta+\mathrm{i}2n\pi\quad(n=0,\pm1,\pm2,\cdots).$$

更一般地，$(z-a)^{\alpha}$（当 α 非正负整数或零时）也有类似情况.

这些例子说明，$(z-a)^{\frac{1}{2}}$，$\ln(z-a)$ 都是**多值函数**，前者具有二值，后者具有无穷个值；类似地，$(z-a)^{-\frac{5}{3}}$ 具有三值，也是多值函数. 另一方面，由它们所代表的右

边每个函数而言，除 $z=a$ 点之外，均为解析函数，称为**多值解析函数的[单值]分支**. 对多值函数，总有某些使两个或更多个分支重合的点，或者使一个分支具有无穷大极限的点，这些点称为**多值函数的支点**，即**多值性奇点**（这里是 $z=a$ 点）. 当变量 z（反时针方向）绕支点（$z=a$）转一圈后，函数（例如 $(z-a)^{1/k}$）从一个分支变为另一分支；若绕过最小的圈数 $k>1$ 之后又回到原先分支，则该点就称为**有限 $k-1$ 阶支点**. 对数 $\ln(z-a)$ 的支点称为**对数支点**，或**无穷阶支点**.

12.2.3 解析函数的 CR 条件 · 共形映射

(1) Cauchy-Riemann 条件

关于解析函数的概念，还可作如下分析.

对实变函数 $f(x)$ 来说，只要其右导数与左导数存在并相同，则 $f(x)$ 就是可求导的，即

$$\lim_{h\to 0^+}\frac{1}{h}[f(x+h)-f(x)]=\lim_{h\to 0^-}\frac{1}{h}[f(x+h)-f(x)]=\frac{\mathrm{d}f}{\mathrm{d}x}.$$

对复变函数 $f(z)$ 来说，则要求在 z 平面上从各方向趋近的极限值存在且相同时，$f(z)$ 才是可求导的，即 $f(z)$ 是解析函数.

由于一个复变函数对应着两个二元实变函数

$$w(z)=w(x+\mathrm{i}z)=u(x,y)+\mathrm{i}v(x,y),$$

因此可从另一角度来进行分析.

$$\left.\frac{\mathrm{d}w}{\mathrm{d}z}\right|_{y不变}=\frac{\partial u}{\partial x}+\mathrm{i}\frac{\partial v}{\partial x}\quad（沿平行于实轴方向趋近时），$$

$$\left.\frac{\mathrm{d}w}{\mathrm{d}z}\right|_{x不变}=\frac{\partial u}{\mathrm{i}\partial y}+\mathrm{i}\frac{\partial v}{\mathrm{i}\partial y}\quad（沿平行于虚轴方向趋近时），$$

而 $\left.\frac{\mathrm{d}w}{\mathrm{d}z}\right|_{y不变}=\left.\frac{\mathrm{d}w}{\mathrm{d}z}\right|_{x不变}=\frac{\mathrm{d}w}{\mathrm{d}z}$ 则要求

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y},\quad \frac{\partial v}{\partial x}=-\frac{\partial u}{\partial y}.$$

这是 $w(z)$ 为解析函数所应满足的条件，称之为 **Cauchy-Riemann 条件**（简称 **CR 条件**），亦称 **Cauchy-Riemann 方程(CR 方程)**.

由 Cauchy-Riemann 方程立即可以得到

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)u=0,$$

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)v=0;$$

即，解析函数的实部及虚部都必须满足二维的 Laplace 方程.

(2) 共形映射

复变函数的解析性还与共形映射的几何概念有着紧密的联系. 如果连续映射 $w=f(z)$在点 z_0 处具有保角性和伸缩不变性,则称该映射为**共形映射**. 显然,若 $f'(z_0)$存在并且非零,则在 $w=f(z)$映射下从 z_0 出发的每个无穷小矢量都伸张 $|f'(z_0)|$倍,都旋转一个相同角度 $\arg f'(z_0)$.

另一方面,一个解析函数 $\zeta=f(z)$,若 $f'(z)$永不为零,还可看作两复变数 $\zeta=\xi+\mathrm{i}\eta$ 与 $z=x+\mathrm{i}y$ 之间的一种坐标变换,称为**共形变换**. 由于 $f(z)$的解析性, Laplace 方程具有不变性,即在此变换下,

$$\nabla^2 u=\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)u=0 \text{ 变为}\left(\frac{\partial^2}{\partial \xi^2}+\frac{\partial^2}{\partial \eta^2}\right)u=0.$$

关于这一点,可证明如下.

$$\frac{\partial}{\partial x}=\left(\frac{\partial \xi}{\partial x}\right)\frac{\partial}{\partial \xi}+\left(\frac{\partial \eta}{\partial x}\right)\frac{\partial}{\partial \eta},$$

$$\frac{\partial^2}{\partial x^2}=\left(\frac{\partial \xi}{\partial x}\right)^2\frac{\partial^2}{\partial \xi^2}+\left(\frac{\partial \eta}{\partial x}\right)^2\frac{\partial^2}{\partial \eta^2}+2\left(\frac{\partial \xi}{\partial x}\right)\left(\frac{\partial \eta}{\partial x}\right)\frac{\partial^2}{\partial \xi\partial \eta}+\left(\frac{\partial^2 \xi}{\partial x^2}\right)\frac{\partial}{\partial \xi}+\left(\frac{\partial^2 \eta}{\partial x^2}\right)\frac{\partial}{\partial \eta};$$

对$\frac{\partial^2}{\partial y^2}$有类似结果. 于是,

$$\begin{aligned}\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}=&\left[\left(\frac{\partial \xi}{\partial x}\right)^2+\left(\frac{\partial \xi}{\partial y}\right)^2\right]\frac{\partial^2}{\partial \xi^2}+\left[\left(\frac{\partial \eta}{\partial x}\right)^2+\left(\frac{\partial \eta}{\partial y}\right)^2\right]\frac{\partial^2}{\partial \eta^2}\\&+2\left[\left(\frac{\partial \xi}{\partial x}\right)\left(\frac{\partial \eta}{\partial x}\right)+\left(\frac{\partial \xi}{\partial y}\right)\left(\frac{\partial \eta}{\partial y}\right)\right]\frac{\partial^2}{\partial \xi\partial \eta}\\&+\left[\left(\frac{\partial^2 \xi}{\partial x^2}\right)+\left(\frac{\partial^2 \xi}{\partial y^2}\right)\right]\frac{\partial}{\partial \xi}+\left[\left(\frac{\partial^2 \eta}{\partial x^2}\right)+\left(\frac{\partial^2 \eta}{\partial y^2}\right)\right]\frac{\partial}{\partial \eta}.\end{aligned}$$

由于 CR 条件

$$\frac{\partial \xi}{\partial x}=\frac{\partial \eta}{\partial y},\quad \frac{\partial \xi}{\partial y}=-\frac{\partial \eta}{\partial x},$$

上式的后两行[ ]中均为零,而首行[ ]中均变为$|f'(z)|^2$;后者

$$\begin{aligned}\left(\frac{\partial \xi}{\partial x}\right)^2+\left(\frac{\partial \xi}{\partial y}\right)^2&=\left(\frac{\partial \xi}{\partial x}\right)^2+\left(\frac{\partial \eta}{\partial x}\right)^2\\&=\left\{\left(\frac{\partial \xi}{\partial x}\right)+\mathrm{i}\left(\frac{\partial \eta}{\partial x}\right)\right\}\left\{\left(\frac{\partial \xi}{\partial x}\right)-\mathrm{i}\left(\frac{\partial \eta}{\partial x}\right)\right\}\\&=\{f'(z)\}\{f'(z)\}^*=|f'(z)|^2.\end{aligned}$$

所以,最后有

$$\nabla^2 u=\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)u=|f'(z)|^2\left(\frac{\partial^2}{\partial \xi^2}+\frac{\partial^2}{\partial \eta}\right)u;$$

若 $f'(z)\neq 0$,则有$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)u=0$ 变为$\left(\frac{\partial^2}{\partial \xi^2}+\frac{\partial^2}{\partial \eta^2}\right)u=0$.

所以,对于具有给定边界的二维 Laplace 方程,可通过共形映射将之变为具有不同边界的 Laplace 方程. 主要目的是这样来选取共形映射,使问题在新边界下易于求解.

12.2.4 解析函数有关定理

下面介绍关于解析函数性质的几个重要定理.

(1) Liouville 定理

定理 若 $f(z)$对 z 平面上所有点都是解析的,而且有界 $|f(z)|<K$,则 $f(z)=\text{const}$.

可以用上节(12.2.3 之(1))的结果来论证这个定理. 对于解析函数

$$f(z) = u(x,y) + \mathrm{i}v(x,y),$$

由 Cauchy-Riemann 条件可知 $u(x,y)$和 $v(x,y)$均满足二维 Laplace 方程

$$\nabla^2 u = 0, \quad \nabla^2 v = 0.$$

想象 $u(x,y)$是(x,y)平面上的势,因为 $f(z)$在平面上处处解析,故处处均应满足无源的势方程$\nabla^2 u=0$,所以,各点的势都是一样的,即 $u(x,y)=\text{const}$. 同理有 $v=\text{const}$. 结果有

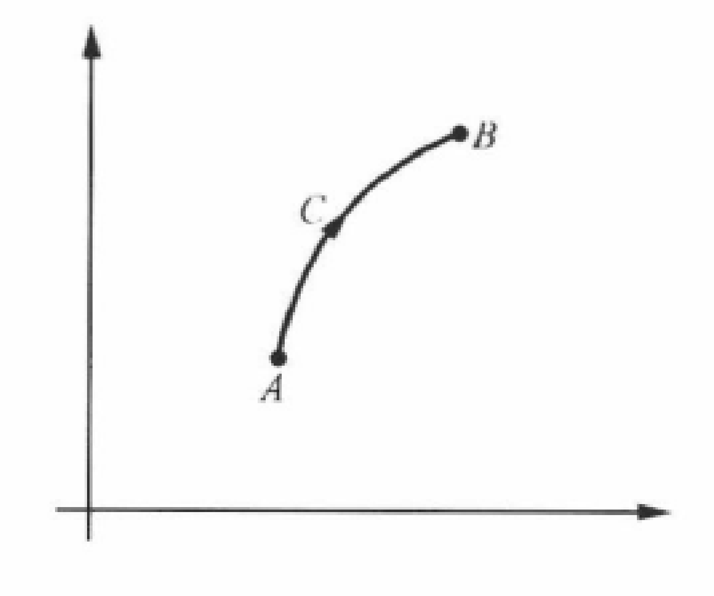
$$f(z) = u(x,y) + \mathrm{i}v(x,y) = \text{const}.$$

这样也算马马虎虎将 Liouville 定理证明了.

(2) 周线积分的 Cauchy 定理

首先介绍复积分的概念. 考虑函数 $f(z)$沿(如下图所示)复平面上一简单曲线 C 从 A 至 B 的**复积分**,记作$\int_C f(z)\mathrm{d}z$. 设曲线 C 可用参数 t 表示为

$$C: \begin{cases} x = x(t), \\ y = y(t), \end{cases} \quad a \leqslant t \leqslant b.$$



于是有

$$\mathrm{d}z = \mathrm{d}(x + \mathrm{i}y) = \left(\frac{\mathrm{d}x}{\mathrm{d}t} + \mathrm{i}\,\frac{\mathrm{d}y}{\mathrm{d}t}\right)\mathrm{d}t,$$

$$f(z) = u(x,y) + \mathrm{i}v(x,y);$$

因此可用实变数 t 的积分来**定义**复积分为

$$\int_C f(z)\mathrm{d}z = \int_a^b (u+\mathrm{i}v)\left(\frac{\mathrm{d}x}{\mathrm{d}t}+\mathrm{i}\frac{\mathrm{d}y}{\mathrm{d}t}\right)\mathrm{d}t = \int_a^b [U(t)+\mathrm{i}V(t)]\mathrm{d}t.$$

所以，沿曲线 C 的复积分是线积分的特例. 对于解析函数，有 $\nabla^2 u=0$，$\nabla^2 v=0$.

例如，

$$\int_{z_a}^{z_b}\mathrm{d}z = \int_a^b\left(\frac{\mathrm{d}x}{\mathrm{d}t}+\mathrm{i}\frac{\mathrm{d}y}{\mathrm{d}t}\right)\mathrm{d}t = (x_b-x_a)+\mathrm{i}(y_b-y_a)$$

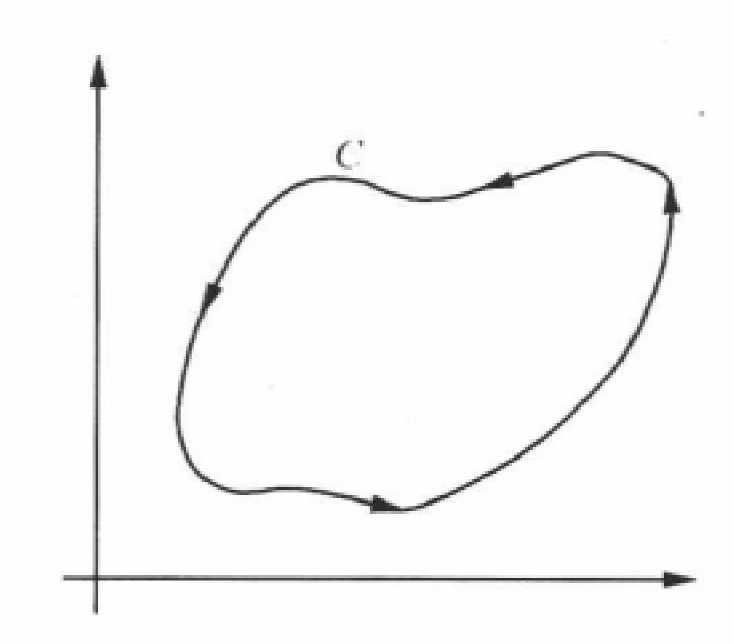
$$= z_b - z_a;$$

$$\int z\mathrm{d}z = \int (x+\mathrm{i}y)\left(\frac{\mathrm{d}x}{\mathrm{d}t}+\mathrm{i}\frac{\mathrm{d}y}{\mathrm{d}t}\right)\mathrm{d}t = \int (x\mathrm{d}x - y\mathrm{d}y)+\mathrm{i}\int (y\mathrm{d}x + x\mathrm{d}y)$$

$$= \frac{1}{2}(x^2-y^2)+\mathrm{i}xy = \frac{1}{2}(x+\mathrm{i}y)^2 = \frac{1}{2}z^2.$$

下面介绍关于**周线积分的 Cauchy 定理**. 若函数 $f(z)$ 在(如下图所示)**周线**(闭合曲线)C 上及 C 内各点都是解析的，则其周线积分为零，即

$$\oint_C f(z)\mathrm{d}z = 0.$$



证 注意到，按复积分定义有

$$\oint_C f(z)\mathrm{d}z = \oint_C (u+\mathrm{i}v)(\mathrm{d}x+\mathrm{i}\mathrm{d}y)$$

$$= \oint_C (u\mathrm{d}x - v\mathrm{d}y) + \mathrm{i}\oint_C (v\mathrm{d}x + u\mathrm{d}y);$$

应用 Green 公式(见第 9 章 9.3.6 节之脚注)

$$\oint_C (P\mathrm{d}x + Q\mathrm{d}y) = \iint_S \left[\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right]\mathrm{d}x\mathrm{d}y,$$

其中 S 是周线 C 所围之面积；于是，上式可化为

$$\oint_C f(z)\mathrm{d}z = -\iint_S\left[\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right]\mathrm{d}x\mathrm{d}y + \mathrm{i}\iint_S\left[\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y}\right]\mathrm{d}x\mathrm{d}y.$$

由于解析函数满足 Cauchy-Riemann 条件，

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y},\quad \frac{\partial v}{\partial x}=-\frac{\partial u}{\partial y},$$

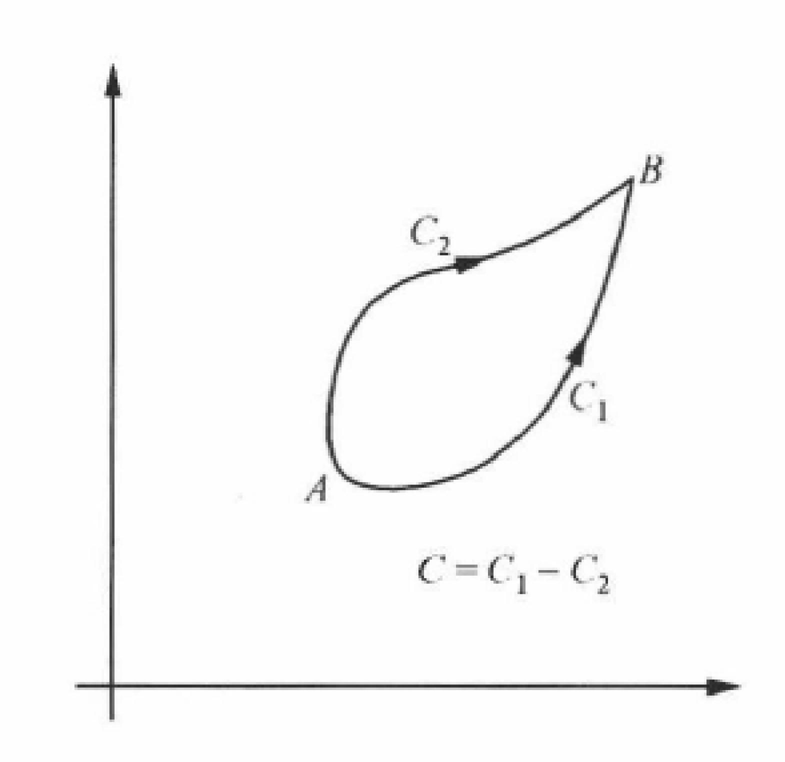
上式右边为零;Cauchy 定理得证.

Cauchy 定理的推论 解析函数的积分值与积分路径无关.

取周线 $C=C_1-C_2$(如下图所示),则有

$$0=\oint_C f(z)\mathrm{d}z=\int_{C_1} f(z)\mathrm{d}z-\int_{C_2} f(z)\mathrm{d}z,$$

推论得证.

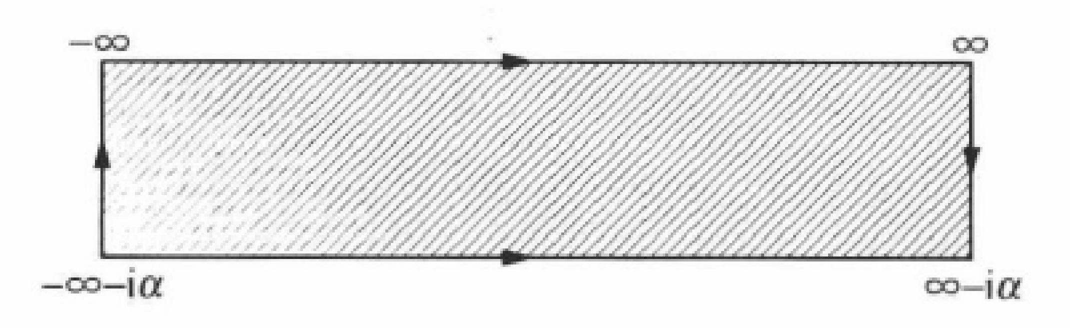


应注意积分路径与积分周线的区别.积分路径是具有端点的曲线,而积分周线则是闭合曲线.由于 Cauchy 定理及其推论,可以适当改变积分路径或使积分周线变形,以便于积分的运算,然而必须注意毋使其触及奇点.

现在应用 Cauchy 定理来论证一下第 8 章 8.4.6 节例 2 中用过的一个结果,

$$\int_{-\infty-\mathrm{i}\alpha}^{\infty-\mathrm{i}\alpha} \mathrm{e}^{-\gamma\xi^2}\mathrm{d}\xi=\int_{-\infty}^{\infty} \mathrm{e}^{-\gamma\xi^2}\mathrm{d}\xi \quad (\gamma>0).$$

将左边复积分 $\int_{C_1} f(z)\mathrm{d}z$ 的积分路径 C_1 从平行于实轴的 $-\infty-\mathrm{i}\alpha$ 至 $\infty-\mathrm{i}\alpha$ 改变为如下图所示的折线 C_2,即由平行于虚轴的两段:$-\infty-\mathrm{i}\alpha$ 至 $-\infty$ 和 ∞ 至 $\infty-\mathrm{i}\alpha$ 以及实轴组成.由于被积函数 $f(z)=\mathrm{e}^{-\gamma z^2}$ 在平行于实轴的任何带内无奇点,故根据 Cauchy 定理的推论有 $\int_{C_1} f(z)\mathrm{d}z=\int_{C_2} f(z)\mathrm{d}z$. 另外,当 $|\mathrm{Re}z|\to\infty$ 时 $f(z)$ 迅速趋于零,平行于虚轴的两段之积分为零.所以,$\int_{C_2} f(z)\mathrm{d}z$ 只剩下沿实轴的 Gauss 积分.



(3) 留数定理

下面来考虑在积分周线内含有函数之奇点的情况. 首先用简单例子来阐明.

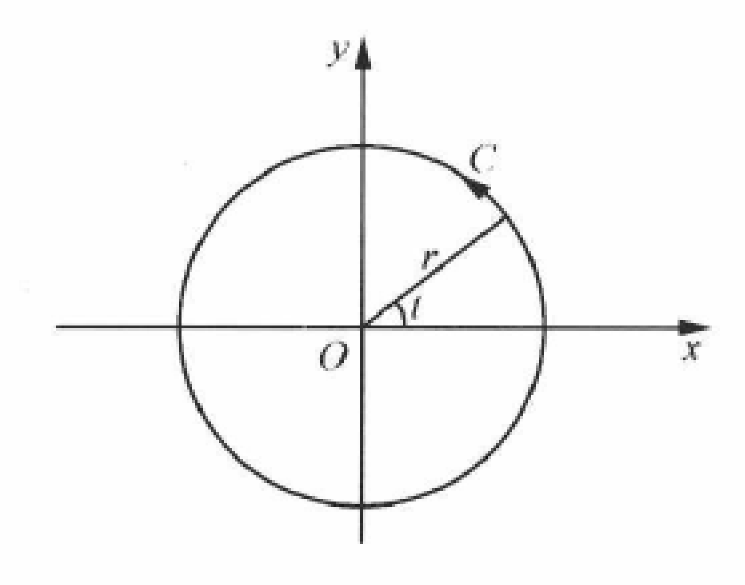
例如，$\oint_C \frac{dz}{z}$，对于$\frac{1}{z}$，$z=0$ 为其单极点. 若所取积分周线 C 为如下图所示中心在 $z=0$ 点以 r 为半径的圆，即

$$C: x = r\cos t, \quad y = r\sin t,$$

$$z = r(\cos t + \mathrm{i}\sin t) = r\mathrm{e}^{\mathrm{i}t},$$

$$\mathrm{d}z = r(-\sin t + \mathrm{i}\cos t)\mathrm{d}t = \mathrm{i}r\mathrm{e}^{\mathrm{i}t}\mathrm{d}t,$$

$$\oint_C \frac{\mathrm{d}z}{z} = \mathrm{i}\int_0^{2\pi}\mathrm{d}t = 2\pi\mathrm{i}.$$



这里约定积分周线 C 取逆时针方向，即，辐角 $\arg z = t$ 增加的方向；否则，上式右边为 $-2\pi\mathrm{i}$.

另一方面，对于$\frac{1}{z^n}$，$z=0$ 为其 n 阶极点.

$$\oint_C \frac{\mathrm{d}z}{z^n} = \mathrm{i}r^{-(n-1)}\int_0^{2\pi}\mathrm{e}^{-\mathrm{i}(n-1)t}\mathrm{d}t = -\frac{r^{-(n-1)}}{(n-1)}\mathrm{e}^{-\mathrm{i}(n-1)t}\Big|_0^{2\pi} = 0.$$

对于一般情况，若函数 $f(z)$ 在 $z=a$ 有 m 阶极点，而在此点周围是解析的，例如，可写成

$$f(z) = \frac{A_{-m}}{(z-a)^m} + \cdots + \frac{A_{-1}}{z-a} + \varphi(z);$$

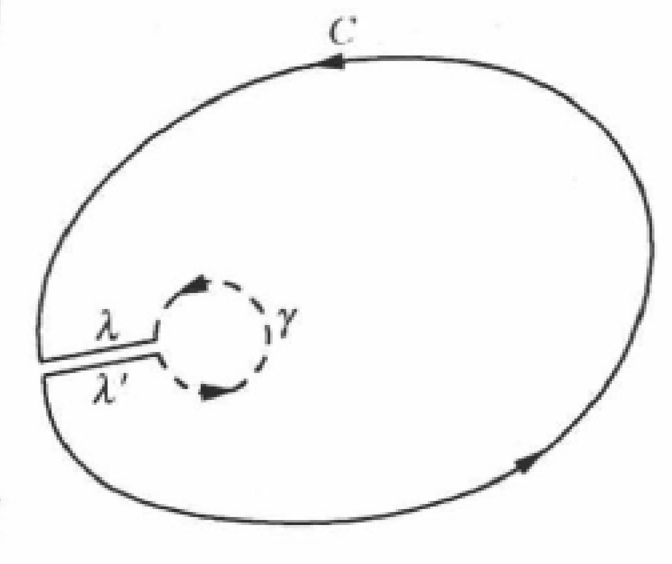
当取积分周线 C 环绕 $z=a$ 点时，显然有

$$\oint_C f(z)\mathrm{d}z = 2\pi\mathrm{i}A_{-1};$$

A_{-1}称为 $f(z)$ 在其孤立奇点 $z=a$ 处的**留数**，常记作 $\mathrm{res}f(a)$.

应该指出，以上结果并不依赖于积分周线 C 的具体形状，只要 $z=a$ 是 C 内唯一奇点. 首先，从前面的阐述可知，上述所有结果均不依赖于圆 γ 的半径 r. 现在取

积分周线 $\Gamma=C+\lambda-\gamma+\lambda'$ 如右图所示，在 Γ 内和 Γ 上均无奇点，可应用 Cauchy 定理得



$$\begin{aligned}0 &= \oint_{\Gamma} f(z)\,\mathrm{d}z \\ &= \int_C f(z)\,\mathrm{d}z - \int_{\gamma} f(z)\,\mathrm{d}z + \int_{\lambda+\lambda'} f(z)\,\mathrm{d}z,\end{aligned}$$

当令圆 γ 的半径 $r\to 0$ 时，割线两边缘之积分值 $\int_{\lambda+\lambda'}$ 相抵消，前两者变为周线积分而有

$$\begin{aligned}\oint_C f(z)\,\mathrm{d}z &= \oint_{\gamma\to 0} f(z)\,\mathrm{d}z \\ &= 2\pi\mathrm{i}\,\mathrm{res}f(a).\end{aligned}$$

上述论证不难推广到周线 C 内含有多个孤立奇点的情况，于是有下列定理.

留数定理 若函数 $f(z)$ 在周线 C 内除有限个孤立奇点 $a_k(k=1,\cdots,m)$ 外是单值解析的，在 C 上无奇点，则

$$\oint_C f(z)\,\mathrm{d}z = 2\pi\mathrm{i}\sum_{k=1}^{m}\mathrm{res}f(a_k).$$

利用留数定理可以很简捷地计算多种类型的积分[①][②]. 现举几个典型例子予以阐明.

(i) $\int_0^{2\pi} R(\sin\theta,\cos\theta)\,\mathrm{d}\theta$，其中 R 是有理函数. 令 $z=\mathrm{e}^{\mathrm{i}\theta}$，取周线 C 为中心在原点的单位圆，容易求得

$$\int_0^{2\pi} R(\sin\theta,\cos\theta)\,\mathrm{d}\theta = \int_C R\left(\frac{z-z^{-1}}{2\mathrm{i}},\frac{z+z^{-1}}{2}\right)\frac{\mathrm{d}z}{\mathrm{i}z} = 2\pi\mathrm{i}\sum_C \mathrm{res},$$

其中 $\sum_C \mathrm{res}$ 代表 $f(z)=\frac{1}{\mathrm{i}z}R\left(\frac{z-z^{-1}}{2\mathrm{i}},\frac{z+z^{-1}}{2}\right)$ 在 C 内留数之和.

(ii) 无穷积分 $\int_{-\infty}^{\infty} Q(x)\,\mathrm{d}x$，其中 $Q(z)$ 在上半平面除有限个孤立奇点外是处处解析的，在实轴没有奇点；在 $0\leqslant\arg z\leqslant\pi$ 中，当 $|z|\to\infty$ 时，一致地有 $zQ(z)\to 0$. 若取周线 C 为沿实轴从 $-R$ 到 R 的直线加上中心在原点、半径为 R 的半圆 C_R（在上半平面，R 充分大），而最后令 $R\to\infty$. 可以证明，在上述条件下，C_R 一段的积分趋于零. 这样一来，容易求得

$$\int_{-\infty}^{\infty} Q(x)\,\mathrm{d}x = 2\pi\mathrm{i}\sum \mathrm{res}Q(z),$$

---

① 较详细介绍可参考：E. T. Whittaker, G. N. Watson, *Modern Analysis*, Ch. Ⅵ.

② 还可参考：郭敦仁,《数学物理方法》,人民教育出版社,北京,1965,第八章.

其中 $\sum \mathrm{res}Q(z)$ 是 $Q(z)$在上半平面的留数之和.

(iii) $\int_0^\infty x^{\alpha-1}Q(x)\mathrm{d}x$ (α 为非整数实数),其中 $Q(z)$在全平面上除有限个孤立奇点外处处是解析的,而且在正实轴上无奇点;当 $|z|\to 0$ 和 $|z|\to\infty$时,$|z^\alpha Q(z)|$一致趋于零,$0\leqslant \arg z\leqslant 2\pi$. 若取周线如图所示,并规定在正实轴上缘 $\arg z=0$,而下缘 $\arg z=2\pi$,则容易求得

$$\int_0^\infty x^{\alpha-1}Q(x)\mathrm{d}x = \frac{2\pi\mathrm{i}}{1-\mathrm{e}^{\alpha 2\pi\mathrm{i}}}\sum \mathrm{res}\{z^{\alpha-1}Q(z)\} \quad (0<\arg z<2\pi)$$

$$= \frac{\pi}{\sin\alpha\pi}\sum \mathrm{res}\{(-z)^{\alpha-1}Q(z)\} \quad (-\pi<\arg(-z)<\pi).$$

其中 $\sum \mathrm{res}\{\cdots\}$ 代表{}内函数在 $Q(z)$的奇点处留数之和.

上述关于积分的计算应用了留数,因而需要有**求留数的公式**. 对于极点的情形,很容易给出如下结果.

$f(z)$在单极点 $z=a$ 的留数,公式是

$$\mathrm{res}f(a) = \lim_{z\to a}\{(z-a)f(z)\}.$$

$f(z)$在 m 阶极点 $z=a$ 的留数,公式是

$$\mathrm{res}f(a) = \lim_{z\to a}\left\{\frac{1}{(m-1)!}\frac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}[(z-a)^m f(z)]\right\}.$$

可以看出,单极点留数的公式是后者在 $m=1$ 时的特例.

(4) Cauchy 积分公式与 Taylor 定理

下面要来讨论解析函数的幂级数展开,其中要用到函数在某一点的各阶导数,包括在该点的函数值.

上一章(11.3.2 节之(2))曾指出,若 v 满足 Laplace 方程 $\nabla^2 v=0$,则 a 点的 v 值可由环绕 a 点的边界值给出,即

$$v_a = \frac{1}{4\pi}\int_\Sigma\left\{\frac{1}{r}\frac{\partial v}{\partial n} - v\frac{\partial}{\partial n}\left(\frac{1}{r}\right)\right\}\mathrm{d}\sigma.$$

一个解析函数 $f(z)$可表达为 $f(z)=u+\mathrm{i}v$,并且满足二维 Laplace 方程 $\nabla^2 u=0$ 和 $\nabla^2 v=0$. 推想可知,$f(z)$在 a 点之值 $f(a)$应可由其环绕 a 点周围之值所决定. 的确有这样的结果,称之为 **Cauchy 积分公式**,可表述如下:设 $f(z)$在周线 C 上及 C 内各点都是解析的,a 为 C 内任意一点,则

$$f(a) = \frac{1}{2\pi\mathrm{i}}\oint_C \frac{f(z)}{z-a}\mathrm{d}z.$$

此公式可用留数定理予以证明. 因为 $z=a$ 是 $\varphi(z)=\dfrac{f(z)}{z-a}$在 C 内的单极点,

$\text{res}\varphi(a)=f(a)$，于是

$$\oint_C \frac{f(z)}{z-a}\mathrm{d}z=\oint\varphi(z)\mathrm{d}z=2\pi\mathrm{i}\,\text{res}\varphi(a)=2\pi\mathrm{i}f(a),$$

公式得证.

若将 Cauchy 积分公式改写为

$$f(z)=\frac{1}{2\pi\mathrm{i}}\oint_C\frac{f(\zeta)}{(\zeta-z)}\mathrm{d}\zeta,$$

则可求得

$$f^{(n)}(z)=\frac{n!}{2\pi\mathrm{i}}\oint_C\frac{f(\zeta)}{(\zeta-z)^{n+1}}\mathrm{d}\zeta.$$

也就是说，有以下关于解析函数各阶导数的定理.

定理 若 $f(z)$在周线 C 内及 C 上是一解析函数，则在 C 内各点其各阶导数 $\frac{\mathrm{d}f}{\mathrm{d}z},\frac{\mathrm{d}^2f}{\mathrm{d}z^2},\cdots$（求导多少次都行）都是解析函数，并可用上面公式表示.

特别是，若有了解析函数 $f(z)$在 a 点的各阶导数值：$f(a),f'(a),f''(a),\cdots$，则有下述定理.

Taylor 定理 解析函数 $f(z)$在 a 点附近可展成 Taylor 级数，它一定收敛，即

$$\begin{aligned}f(z)&=f(a)+f'(a)(z-a)+\frac{1}{2!}f''(a)(z-a)^2+\cdots\\&=\sum_{n=0}^{\infty}\frac{1}{n!}f^{(n)}(a)(z-a)^n;\end{aligned}$$

若 $f(z)$在以 a 为中心、R 为半径的圆 C 上及其内无任何奇点，则

$$f^{(n)}(a)=\frac{n!}{2\pi\mathrm{i}}\oint_C\frac{f(z)}{(z-a)^{n+1}}\mathrm{d}z.$$

注意到，将 $f(z)=\frac{1}{2\pi\mathrm{i}}\oint_C\frac{f(\zeta)}{\zeta-z}\mathrm{d}\zeta$ 中的$\frac{1}{\zeta-z}$ 作如下展开，

$$\begin{aligned}\frac{1}{\zeta-z}&=\frac{1}{(\zeta-a)-(z-a)}\\&=\frac{1}{(\zeta-a)}+\frac{(z-a)}{(\zeta-a)}\cdot\frac{1}{(\zeta-a)-(z-a)}\\&=\frac{1}{\zeta-a}+\frac{(z-a)}{(\zeta-a)^2}+\frac{(z-a)^2}{(\zeta-a)^2}\cdot\frac{1}{(\zeta-a)-(z-a)}\\&=\cdots\\&=\frac{1}{\zeta-a}+\frac{z-a}{(\zeta-a)^2}+\frac{(z-a)^2}{(\zeta-a)^3}+\cdots+\frac{(z-a)^n}{(\zeta-a)^{n+1}}+\frac{(z-a)^{n+1}}{(\zeta-a)^{n+1}}\cdot\frac{1}{\zeta-z},\end{aligned}$$

就很容易得到上述 Taylor 级数展开式，并可证明此级数在圆 C 内收敛.

12.2.5　解析函数的表示方法与解析延拓

解析函数可用无穷级数表示或用积分表示，偶尔还有用无穷乘积表示，下面将对前两种方法分别作进一步阐述.

(1) 无穷级数表示.收敛半径

设无穷级数 $\sum_{n=0}^{\infty} f_n(z)$ 沿周线 C 一致收敛，而 $f_n(z)$ 在 C 上和 C 内处处解析；则 $\sum_{n=0}^{\infty} f_n(z)$ 收敛，此级数之和是在 C 上和 C 内处处解析的函数.

具体而且重要的是幂级数，如前面讨论过的 Taylor 级数，其中 $f_n(z)=c_n(z-a)^n$. 将一函数展开成幂级数时，其**收敛半径**就等于从展开的一点到最近的奇点之间的距离；以此半径所作之圆称为**收敛圆**；级数在此圆内收敛，圆上必有一点是函数的奇点.

例如：对实变数，$1-x+x^2-x^3+\cdots$ 的收敛区间是 $|x|<1$ 或开区间 $(-1,1)$，而对复变数，$1-z+z^2-z^3+\cdots$ 的收敛圆是 $|z|<1$，收敛半径为 1（因为 $z=-1$ 是奇点）；它是 $\frac{1}{1+z}$ 在 $z=0$ 点的展开. 应注意，在此收敛圆内才有

$$\frac{1}{1+z}=1-z+z^2-z^3+\cdots=\sum_{n=0}^{\infty}(-1)^n z^n,\quad |z|<1.$$

也就是说，一个函数当展成级数时，只有在其收敛圆内，此级数才是该函数的表示；在收敛圆外则不然，要用另外的级数来表示. 所以，级数和它所代表的函数要分开，这观念很重要！

还以 $\frac{1}{1+z}$ 为例，当 $|z|>1$ 时要用另一级数表示，即在 $\frac{1}{z}=0$（即 $z=\infty$）点，可展开为

$$\frac{1}{1+z}=\frac{1}{z\left(1+\frac{1}{z}\right)}=\frac{1}{z}-\frac{1}{z^2}+\frac{1}{z^3}-\frac{1}{z^4}+\cdots,$$

$$|z|>1\ \left(\left|\frac{1}{z}\right|<1\right).$$

所以，$\frac{1}{1+z}$ 在复平面上 $|z|=1$ 之内外由两部分级数表示，即

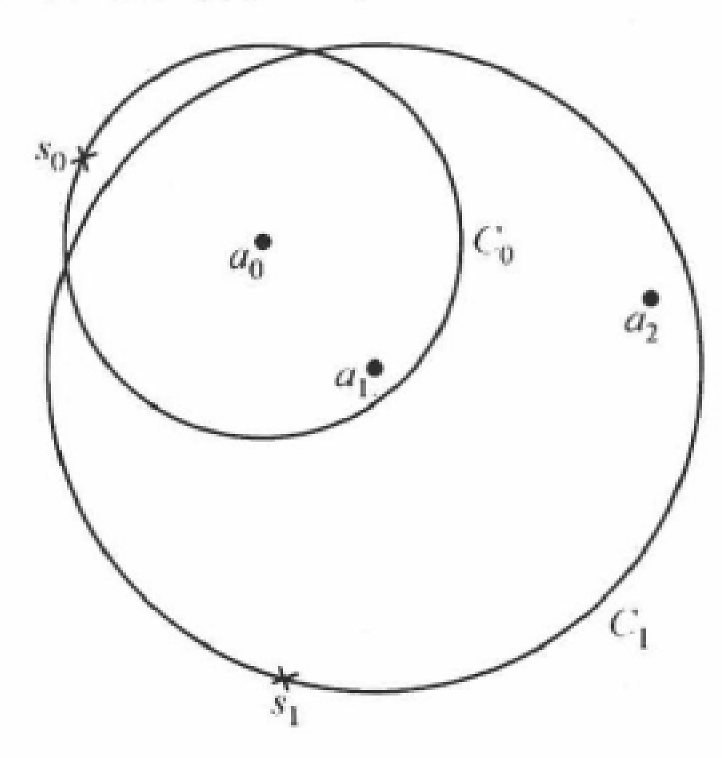
$$\frac{1}{1+z}=1-z+z^2-z^3+\cdots=\sum_{n=0}^{\infty}(-)^n z^n,\quad |z|<1,$$

$$\frac{1}{1+z}=\frac{1}{z}-\frac{1}{z^2}+\frac{1}{z^3}-\frac{1}{z^4}+\cdots=\sum_{n=0}^{\infty}(-)^n\frac{1}{z^{n+1}},\quad |z|>1.$$

(2) 解析延拓. 解析函数的 Weierstrass 定义

一个函数在不同区域内可以有不同的级数表示，如何去辨识它呢？这就需要首先引进解析延拓的概念.

若已知 $f(z)$在 a_0 点邻域可展成 Taylor 级数(参考下图)：



$$f_0(z)=\sum c_n(z-a_0)^n,$$

其收敛圆为 C_0，C_0 上 s_0 点为奇点. 在 C_0 内选一点 a_1，$f_0(z)$在此点是解析的，可求出各阶导数 $f_0(a_1)$，$f_0'(a_1)$，…；又可在 a_1 点邻域展成 Taylor 级数：

$$f_1(z)=\sum c_n^{(1)}(z-a_1)^n,$$

其收敛圆为 C_1，C_1 上 s_1 点为奇点. 圆 C_1 除了有一部分与圆 C_0 重叠外，还会包括圆 C_0 以外的一部分. 再在 C_1 内选一点 a_2，又可求得在收敛圆 C_2 内的级数表示 $f_2(z)$；又会延拓出新范围. 这样继续进行下去，就可将复平面上除奇点以外的部分全部包括在内. 这种过程就是所谓 Weierstrass **解析延拓**方法.

因此，从一个起始幂级数出发，加上经解析延拓所可能得到的一切幂级数的全体，构成由起始幂级数生成的(Weierstrass 意义下的)**完全解析函数**.

根据 Weierstrass 关于解析函数的上述定义，立即可得关于两函数同一性的 **Weierstrass 定理**：两个不同的解析表达式，若它们所代表的幂级数彼此能通过解析延拓而获得，则它们定义同一函数.

(3) 积分表示

解析函数还可以用含参数的积分来表示，可由一般公式

$$f(z)=\int_C K(z,\zeta)\varphi(\zeta)\mathrm{d}\zeta$$

描述；其中 $K(z,\zeta)$是积分表示的核，$\varphi(\zeta)$是它的密度，C 是积分周线(或积分路径). 当然，也要求积分一致收敛，即，需同时给出其收敛条件.

例如，Γ **函数**可定义为

$$\Gamma(z)=\int_0^\infty t^{z-1}e^{-t}dt,\quad \mathrm{Re}z>0,$$

它定义了右半平面内的 $\Gamma(z)$值.也就是说,积分表示跟级数表示一样,也是在某一区域内的表示.所以,同样存在解析延拓的问题.通过分部积分容易证明,$\Gamma(z)$满足下列关系

$$\Gamma(z+1)=z\Gamma(z),$$

用这个关系可将 $\Gamma(z)$延拓到负的区域去,但是,$z=0,-1,-2,\cdots$是极点.如果取积分周线如下图所示$\left(记作\int_\infty^{(0+)}\right)$,则得到积分表示为

$$\Gamma(z)=\frac{1}{e^{2\pi iz}-1}\int_\infty^{(0+)}\zeta^{z-1}e^{-\zeta}d\zeta\quad(0\leqslant\arg z\leqslant 2\pi),$$



它对除点 $z=n$(整数)之外的所有 z 成立,而在这些点处 $\Gamma(z)$具有单极点.

顺便说一句,Cauchy 积分公式

$$f(z)=\frac{1}{2\pi i}\int_C\frac{f(\zeta)}{\zeta-z}d\zeta$$

用周线 C 上的值 $f(\zeta)$来表示 C 内的值,也是一种积分表示.

还应指出,对积分表示的核 $K(z,\zeta)$,最常用的是 $e^{iz\zeta}$(Fourier 核),$e^{-z\zeta}$(Laplace 核),ζ^{z-1}(Mellin 核)等.由此可以看出,解析函数的积分表示与积分变换方法(见 8.4,8.6 节)密切相关.[对后者而言,$K(z,\zeta)$称为积分变换的核.]

12.2.6 Γ 函数和 B 函数①

下面顺便对 Γ 函数作进一步讨论.

再次写出 Γ 函数的定义

$$\Gamma(z)=\int_0^\infty e^{-t}t^{z-1}dt,\quad \mathrm{Re}z>0.$$

(1) 递推关系

$\Gamma(z)$满足下列**递推关系**

$$\Gamma(z+1)=z\Gamma(z),$$

反复应用此关系 n 次可得

$$\Gamma(z+n)=(z+n-1)(z+n-2)\cdots(z+1)z\Gamma(z),$$

因而可将 $\Gamma(z)$的定义推广为

① 王竹溪,郭敦仁:《特殊函数概论》,北京大学出版社,北京,2000;3.1,3.2,3.7,3.8 诸节.

$$\Gamma(z)=\frac{\Gamma(z+n)}{(z)_n}=\frac{1}{(z)_n}\int_0^\infty e^{-t}t^{z+n-1}dt,\quad \mathrm{Re}z>-n;$$

n 为任意正整数，而

$$(z)_n=\frac{\Gamma(z+n)}{\Gamma(z)}=z(z+1)\cdots(z+n-1).$$

由定义和递推关系很容易得出

$$\Gamma(1)=0!=1,\quad \Gamma(n+1)=n!.$$

另外，$\Gamma(z)$在有限区域内的奇点都是一阶极点，极点为 $z=0,-1,-2,\cdots,-n,\cdots$，在极点 $z=-n$ 处的留数是

$$\lim_{z\to-n}(z+n)\Gamma(z)=\left.\frac{\Gamma(z+n+1)}{z(z+1)\cdots(z+n-1)}\right|_{z=-n}=\frac{(-)^n}{n!}.$$

(2) B 函数

首先，考虑两个 Γ 函数的乘积

$$\Gamma(p)\Gamma(q)=\int_0^\infty e^{-u}u^{p-1}du\int_0^\infty e^{-v}v^{q-1}dv.$$

令 $u=x^2,v=y^2$，得

$$\Gamma(p)\Gamma(q)=4\int_0^\infty\int_0^\infty e^{-(x^2+y^2)}x^{2p-1}y^{2q-1}dxdy.$$

再令 $x=r\cos\theta,y=r\sin\theta$，得

$$\Gamma(p)\Gamma(q)=4\int_0^\infty e^{-r^2}r^{2(p+q)-1}dr\int_0^{\frac{\pi}{2}}(\cos\theta)^{2p-1}(\sin\theta)^{2q-1}d\theta.$$

前一积分中令 $r^2=t$，后一积分中令 $\cos^2\theta=x$，分别得

$$2\int_0^\infty e^{-r^2}r^{2(p+q)-1}dr=\int_0^\infty e^{-t}t^{p+q-1}dt=\Gamma(p+q),$$

$$2\int_0^{\frac{\pi}{2}}(\cos\theta)^{2p-1}(\sin\theta)^{2q-1}d\theta=\int_0^1 x^{p-1}(1-x)^{q-1}dx.$$

现在，定义 **B 函数**为

$$B(p,q)=\int_0^1 x^{p-1}(1-x)^{q-1}dx,\quad \mathrm{Re}p>0,\mathrm{Re}q>0.$$

于是，上面显然证明了有

$$B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}=B(q,p),$$

此式可作为 B 函数的推广定义，它不受上述条件 $\mathrm{Re}p>0,\mathrm{Re}q>0$ 的限制.

对于后一积分，还可通过令 $x=t/(1+t)$而变为

$$B(p,q)=\int_0^\infty\frac{t^{p-1}dt}{(1+t)^{p+q}}=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

当 $p+q=1$ 而 $0<\mathrm{Re}p<1$ 时，此积分可用留数定理求出，按照12.2.4节(3)之例 3

中给出的公式$[x\to t,\alpha\to p,Q(z)=(1+z)^{-1}]$

$$\int_0^{\infty}\frac{t^{p-1}}{1+t}\mathrm{d}t=\frac{\pi}{\sin p\pi}\sum\mathrm{res}\{(-z)^{p-1}(1+z)^{-1}\}=\frac{\pi}{\sin p\pi}.$$

$$(-\pi<\arg(-z)<\pi)$$

因此得到 Γ 函数的下列重要性质($p\to z$)：

$$\Gamma(z)\Gamma(1-z)=\frac{\pi}{\sin\pi z}.$$

此式可不受 $0<\mathrm{Re}z<1$ 的限制，因为左右两边都是除了 $z=n$(整数)之外全平面的解析函数. 由此式的一个特例 $z=\frac{1}{2}$可得

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^2=\pi,\quad \Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}.$$

另外，在B函数定义之前的二式中分别令 $2(p+q)-1=\lambda$ 和 $2p-1=\alpha,2q-1=\beta$，可得下列积分公式：

$$\int_0^{\infty}\mathrm{e}^{-r^2}r^{\lambda}\mathrm{d}r=\frac{1}{2}\Gamma\left(\frac{\lambda+1}{2}\right)\quad 及$$

$$\int_0^{\infty}\mathrm{e}^{-r^2}\mathrm{d}r=\frac{1}{2}\Gamma\left(\frac{1}{2}\right)=\frac{1}{2}\sqrt{\pi};$$

$$\int_0^{\frac{\pi}{2}}(\cos\theta)^{\alpha}(\sin\theta)^{\beta}\mathrm{d}\theta=\frac{\Gamma\left(\frac{\alpha+1}{2}\right)\Gamma\left(\frac{\beta+1}{2}\right)}{2\Gamma\left(\frac{\alpha+\beta}{2}+1\right)}.$$

(3) 周线积分表示

Γ 函数的积分表示

$$\Gamma(z)=\int_0^{\infty}\mathrm{e}^{-t}t^{z-1}\mathrm{d}t\quad(\mathrm{Re}z>0)$$

$$=\frac{1}{(z)_n}\int_0^{\infty}\mathrm{e}^{-t}t^{z+n-1}\mathrm{d}t\quad(\mathrm{Re}z>-n),$$

前者仅适用于右半平面，后者适用范围有扩展，但极点 $z=0,-1,-2,\cdots,-n$ 除外.

现在考虑下列周线积分(所取周线如下图所示)

$$\int_{\infty}^{(0+)}\mathrm{e}^{-t}t^{z-1}\mathrm{d}t=\int_{\infty}^{\delta}\mathrm{e}^{-t}t^{z-1}\mathrm{d}t+\int_0^{2\pi}\mathrm{e}^{-\delta\mathrm{e}^{\mathrm{i}\theta}}(\delta\mathrm{e}^{\mathrm{i}\theta})^{z-1}\delta\mathrm{e}^{\mathrm{i}\theta}\mathrm{i}\mathrm{d}\theta+\mathrm{e}^{2\pi\mathrm{i}z}\int_{\delta}^{\infty}\mathrm{e}^{-t}t^{z-1}\mathrm{d}t,$$

最后一个积分前的因子是由于正向绕原点一周后 $\arg t=2\pi$. 若先限制 $\mathrm{Re}z>0$，则当 $\delta\to0$ 时，右边对 θ 的积分项含 δ^2 因子而为零. 结果容易得到$\Gamma(z)$的周线积分

表示：

$$\Gamma(z)=\frac{1}{e^{2\pi i z}-1}\int_{\infty}^{(0+)}e^{-t}t^{z-1}dt\quad(|\arg t|<2\pi)$$

$$=-\frac{1}{2i\sin\pi z}\int_{\infty}^{(0+)}e^{-t}(-t)^{z-1}dt\quad(|\arg(-t)|<\pi),$$

它们适用于除 z 为整数之外的所有 z 值.

利用上段中的

$$\Gamma(z)\Gamma(1-z)=\frac{\pi}{\sin\pi z},$$

可得下列表达式

$$\frac{1}{\Gamma(1-z)}=-\frac{1}{2\pi i}\int_{\infty}^{(0+)}e^{-t}(-t)^{z-1}dt\quad(|\arg(-t)|<\pi).$$

它适用于任意 z 值，包括 z 等于整数.

在上式中令 $1-z\to z$，可得

$$\frac{1}{\Gamma(z)}=-\frac{1}{2\pi i}\int_{\infty}^{(0+)}e^{-t}(-t)^{-z}dt\quad(|\arg(-t)|<\pi);$$

或者再令 $t\to -t$，可得

$$\frac{1}{\Gamma(z)}=\frac{1}{2\pi i}\int_{-\infty}^{(0+)}e^{t}t^{-z}dt\quad(|\arg t|<\pi);$$

后者取周线如下图所示. 最后二式对任何 z 值均适用，可作为$\Gamma(z)$的普遍表达式.

$-\infty$

12.3 常点邻域内的级数解

从本节开始，将要讨论二阶线性常微分方程

$$\frac{d^2u}{dz^2}+p(z)\frac{du}{dz}+q(z)u=0$$

的具体解法.

12.3.1 方程的奇点与常点

方程系数 $p(z)$或 $q(z)$的每个奇点都是**方程的奇点**，因为它们的总体将构成方程通解中的全部奇点和支点. 另外，$p(z)$和 $q(z)$都是解析函数的其余一切 z 值，则称为**方程的常点**.

至于**无穷远点** $z=\infty$，需另行考虑. 通过作变换 $z_1=\dfrac{1}{z}$，

$$\frac{\mathrm{d}u}{\mathrm{d}z}=\frac{\mathrm{d}z_1}{\mathrm{d}z}\frac{\mathrm{d}u}{\mathrm{d}z_1}=-\frac{1}{z^2}\frac{\mathrm{d}u}{\mathrm{d}z_1}=-z_1^2\frac{\mathrm{d}u}{\mathrm{d}z_1},$$

$$\frac{\mathrm{d}^2u}{\mathrm{d}z^2}=-z_1^2\frac{\mathrm{d}}{\mathrm{d}z_1}\left(-z_1^2\frac{\mathrm{d}u}{\mathrm{d}z_1}\right)=z_1^4\frac{\mathrm{d}^2u}{\mathrm{d}z_1^2}+2z_1^3\frac{\mathrm{d}u}{\mathrm{d}z_1},$$

上述微分方程变换为

$$\frac{\mathrm{d}^2u}{\mathrm{d}z_1^2}+\left[\frac{2}{z_1}-\frac{1}{z_1^2}p\left(\frac{1}{z_1}\right)\right]\frac{\mathrm{d}u}{\mathrm{d}z_1}+\frac{1}{z_1^4}q\left(\frac{1}{z_1}\right)u=0;$$

因而要看 $z_1=0$ 是否为此方程的奇点,即是否为

$$p_1(z_1)=\frac{2}{z_1}-\frac{1}{z_1^2}p\left(\frac{1}{z_1}\right)\quad 和\quad q_1(z_1)=\frac{1}{z_1^4}q\left(\frac{1}{z_1}\right)$$

两者之一的奇点.

例如,**超几何微分方程**

$$z(1-z)\frac{\mathrm{d}^2u}{\mathrm{d}z^2}+\{\gamma-(\alpha+\beta+1)z\}\frac{\mathrm{d}u}{\mathrm{d}z}-\alpha\beta u=0,$$

对此有

$$p(z)=\frac{\gamma-(\alpha+\beta+1)z}{z(1-z)},\quad q(z)=\frac{\alpha\beta}{z(z-1)};$$

显然,$z=0$ 和 $z=1$ 是 p 与 q 的极点,因而是方程的奇点. 至于 $z=\infty$点,由于

$$\frac{1}{z_1^4}q\left(\frac{1}{z_1}\right)=\frac{1}{z_1^4}\frac{\alpha\beta}{\frac{1}{z_1}\left(\frac{1}{z_1}-1\right)}=\frac{\alpha\beta}{z_1^2(1-z_1)},$$

$z_1=0$ 是其极点;即 $z=\infty$为原方程的奇点.

对于方程的常点与奇点,需采用不同的求解方法. 奇点邻域内的解法将在以后两节再来介绍,本节将专门讨论常点邻域内的 **Taylor 级数解法**.

若 $p(z)$和 $q(z)$在 $z=z_0$ 点是解析的,当给出 u 和$\frac{\mathrm{d}u}{\mathrm{d}z}$在该点的初值后,就可用微分方程求得任何高阶导数$\frac{\mathrm{d}^mu}{\mathrm{d}z^m}(m=2,3,\cdots)$在该点之值;因此得到该点邻域的 Taylor 级数. 它是在其收敛圆内方程的解. 然后,再通过解析延拓,可将此解扩展至求解区域内除奇点以外的全部范围.

12.3.2 Legendre 微分方程

下面将以 **Legendre 微分方程**

$$(1-z^2)\frac{\mathrm{d}^2u}{\mathrm{d}z^2}-2z\frac{\mathrm{d}u}{\mathrm{d}z}+\lambda(\lambda+1)u=0$$

作为具体例子,进一步阐明求级数解的步骤.

因为

$$p(z)=-\frac{2z}{1-z^2},\quad q(z)=\frac{\lambda(\lambda+1)}{(1-z^2)},$$

$z=\pm1$ 是方程的奇点. 对于 $z=\infty$，令 $z_1=\frac{1}{z}$ 后得到

$$p_1(z_1)=\frac{2}{z_1}-\frac{1}{z_1^2}p\left(\frac{1}{z_1}\right)=\frac{2z_1}{z_1^2-1},$$
$$q_1(z_1)=\frac{1}{z_1^4}q\left(\frac{1}{z_1}\right)=\frac{\lambda(\lambda+1)}{z_1^2(z_1^2-1)};$$

由此可知 $z_1=0$ 是 $q_1(z_1)$ 的二阶极点，即 $z=\infty$ 亦为方程的奇点. 除 $z=\pm1,\infty$ 之外，其余都是方程的常点.

现在来求 Legendre 方程在其常点 $z=0$ 邻域的 **Taylor 级数解**，给定初值 $u(0)=c_1,\left.\frac{\mathrm{d}u}{\mathrm{d}z}\right|_{z=0}=c_2$.

显然，在常点 $z=0$ 邻域，方程的解为 z 的幂级数，可设为

$$u=\sum_{n=0}^{\infty}a_nz^n.$$

将 $u=\sum_{n=0}^{\infty}a_nz^n$ 代入 Legendre 方程后得到

$$(1-z^2)\sum_{n=0}^{\infty}n(n-1)a_nz^{n-2}-2z\sum_{n=0}^{\infty}na_nz^{n-1}+\lambda(\lambda+1)\sum_{n=0}^{\infty}a_nz^n=0,$$

分别将 $(1-z^2)$ 和 z 乘进去后，此方程在求和号下只出现 z^n 和 z^{n-2} 这样的幂，可将其写成

$$\sum_{n=0}^{\infty}n(n-1)a_nz^{n-2}-\sum_{n=0}^{\infty}n(n-1)a_nz^n-2\sum_{n=0}^{\infty}na_nz^n+\lambda(\lambda+1)\sum_{n=0}^{\infty}a_nz^n=0,$$

注意到第一项中 $n=0$ 和 $n=1$ 时系数为零，故可将 $n\to n+2$ 而得到

$$\sum_{n=0}^{\infty}\{(n+2)(n+1)a_{n+2}-[n(n-1)+2n-\lambda(\lambda+1)]a_n\}z^n=0.$$

由于 z 为变量，必须其各次幂的系数均为零，上式才得以成立. 所以有

$$(n+1)(n+2)a_{n+2}-[n^2+n-\lambda(\lambda+1)]a_n=0,$$

即，得到下列**递推关系**

$$a_{n+2}=\frac{n(n+1)-\lambda(\lambda+1)}{(n+1)(n+2)}a_n=\frac{(n-\lambda)(n+\lambda+1)}{(n+1)(n+2)}a_n.$$

因为由初值条件 $u(0)=c_1,\left.\frac{\mathrm{d}u}{\mathrm{d}z}\right|_{z=0}=c_2$ 可得

$$u(0)=\left(\sum a^nz^n\right)_{z=0}=a_0=c_1,$$
$$\left.\frac{\mathrm{d}u}{\mathrm{d}z}\right|_{z=0}=\left.\left(\sum na_nz^{n-1}\right)\right|_{z=0}=a_1=c_2.$$

于是,利用递推关系,由 $a_0=c_1$ 可决定所有偶次幂的系数 a_{2n},由 $a_1=c_2$ 可决定所有奇次幂的系数 a_{2n+1}. 结果如下(以 $\lambda=\frac{1}{2}$ 为例)

$$a_{2n}=\frac{(4n-1)(4n-5)}{2^2(2n)(2n-1)}a_{2(n-1)}$$
$$a_0=c_1$$
$$a_2=\frac{3(-1)}{2^2\cdot 2\cdot 1}c_1=-\frac{3}{8}c_1$$
$$a_4=\frac{7\cdot 3\cdot 3\cdot(-1)}{2^4(4!)}c_1=-\frac{21}{128}c_1$$
$$\vdots$$
$$a_{2n}=-\frac{(4n-1)\prod_{k=0}^{n-1}(4k-1)^2}{2^{2n}(2n)!}c_1$$

$$a_{2n+1}=\frac{(4n+1)(4n-3)}{2^2(2n+1)(2n)}a_{2n-1}$$
$$a_1=c_2$$
$$a_3=\frac{5\cdot 1}{2^2\cdot 3\cdot 2}c_2=\frac{5}{24}c_2$$
$$a_5=\frac{9\cdot 5\cdot 5\cdot 1}{2^4(5!)}c_2=\frac{15}{128}c_2$$
$$\vdots$$
$$a_{2n+1}=\frac{(4n+1)\prod_{k=0}^{n-1}(4k+1)^2}{2^{2n}(2n+1)!}c_2$$

最后求得 $z=0$ 邻域的幂级数解为

$$\begin{aligned}u&=c_1\left\{1-\frac{3}{8}z^2-\frac{21}{128}z^4-\cdots\right\}+c_2\left\{z+\frac{5}{24}z^3+\frac{15}{128}z^5+\cdots\right\}\\&=c_1u_1+c_2u_2.\end{aligned}$$

容易看出,在 $z=0$ 有

$$u_1(0)=1,\quad \left.\frac{\mathrm{d}u_1}{\mathrm{d}z}\right|_{z=0}=0;$$
$$u_2(0)=0,\quad \left.\frac{\mathrm{d}u_2}{\mathrm{d}z}\right|_{z=0}=1.$$

下面用**检比法**来检验级数的收敛性. 对于无穷级数 $\sum_{n=0}^{\infty}f_n(z)$,若

$$\lim_{n\to\infty}\left|\frac{f_{n+1}(z)}{f_n(z)}\right|<1,$$

则级数收敛. 对目前的情况,可分别讨论 u_1 和 u_2 的收敛性;

$$u_1:\ \lim_{n\to\infty}\left|\frac{a_{2n}}{a_{2n-2}}z^2\right|=\lim_{n\to\infty}\left|\frac{(4n-1)(4n-5)}{2^2(2n)(2n-1)}z^2\right|=|z^2|,$$
$$u_2:\ \lim_{n\to\infty}\left|\frac{a_{2n+1}}{a_{2n-1}}z^2\right|=\lim_{n\to\infty}\left|\frac{(4n+1)(4n-3)}{2^2(2n+1)(2n)}z^2\right|=|z^2|.$$

所以,两部分级数的收敛条件都是 $|z^2|<1$,即 $|z|<1$;因此,上述幂级数 u 在单位圆内是收敛的. 此结果亦可从方程的奇点 $z=\pm 1$ 直接得到.

下面再举一个简单例子,求

$$\frac{\mathrm{d}^2u}{\mathrm{d}z^2}+u=0$$

在 $z=0$ 的幂级数解. 设 $u=\sum\limits_{n=0}^{\infty}a_nz^n$，容易得到

$$\sum_{n=0}^{\infty}(n+2)(n+1)a_{n+2}z^n+\sum_{n=0}^{\infty}a_nz^n=0,$$

$$a_{n+2}=-\frac{1}{(n+2)(n+1)}a_n,$$

$$u=c_1\left\{1-\frac{1}{2!}z^2+\frac{1}{4!}z^4-\frac{1}{6!}z^6+\cdots\right\}+c_2\left\{z-\frac{1}{3!}z^3+\frac{1}{5!}z^5-\frac{1}{7!}z^7+\cdots\right\}$$

$$=c_1\cos z+c_2\sin z.$$

由于方程的唯一奇点是 $z=\infty$，上面的级数分别收敛于 $\cos z$ 和 $\sin z$[①]. 所以，这一结果与本章开头(12.1.1 节)所给一致.

12.3.3 级数解法的具体步骤

现在按照上面求 Legendre 方程级数解的例子，总结一下二阶线性常微分方程

$$\frac{\mathrm{d}^2u}{\mathrm{d}z^2}+p(z)\frac{\mathrm{d}u}{\mathrm{d}z}+q(z)u=0$$

的级数解法，大致可分为以下几个具体步骤：

(i) **标志出方程的奇点**

在复平面上将方程系数 $p(z)$和 $q(z)$的一切奇点全都标志出来，它们全是方程的奇点；分析无穷远点的情形，判定其是否为方程的奇点.

(ii) **选定起始常点及幂级数**

选定方程的一个常点 z_0 作为起点，则在此点邻域内的幂级数解形式为 $u=f_0(z)=\sum\limits_{n=0}^{\infty}a_n(z-z_0)^n$，其中 a_n 等为待定系数. 还应指出，当 $p(z)$和 $q(z)$并非有理函数[②]时，需要同时将它们展开成$(z-z_0)$的幂级数.

(iii) **确定系数**

将以上幂级数代入微分方程(当 p,q 为分式有理函数时，可先将微分方程通分后再将解代入)，令所得方程中各同次幂系数之和为零，从这一系列方程可得到确定系数 a_n 的递推关系. 解中仍可含两个任意常数.

(iv) **确定收敛半径**

根据奇点分布情况或用检比法来确定此幂级数解的收敛半径.

---

① 由 $e^{iz}=\cos z+i\sin z$ 以及 e^{iz}的幂级数展开式

$$e^{iz}=\sum_{n=0}^{\infty}\frac{1}{n!}(iz)^n=\sum_{m=0}^{\infty}\frac{(-1)^m}{(2m)!}z^{2m}+i\sum_{m=0}^{\infty}\frac{(-1)^m}{(2m+1)!}z^{2m+1},$$

立即可得此结果.

② 两个多项式之比为有理函数.

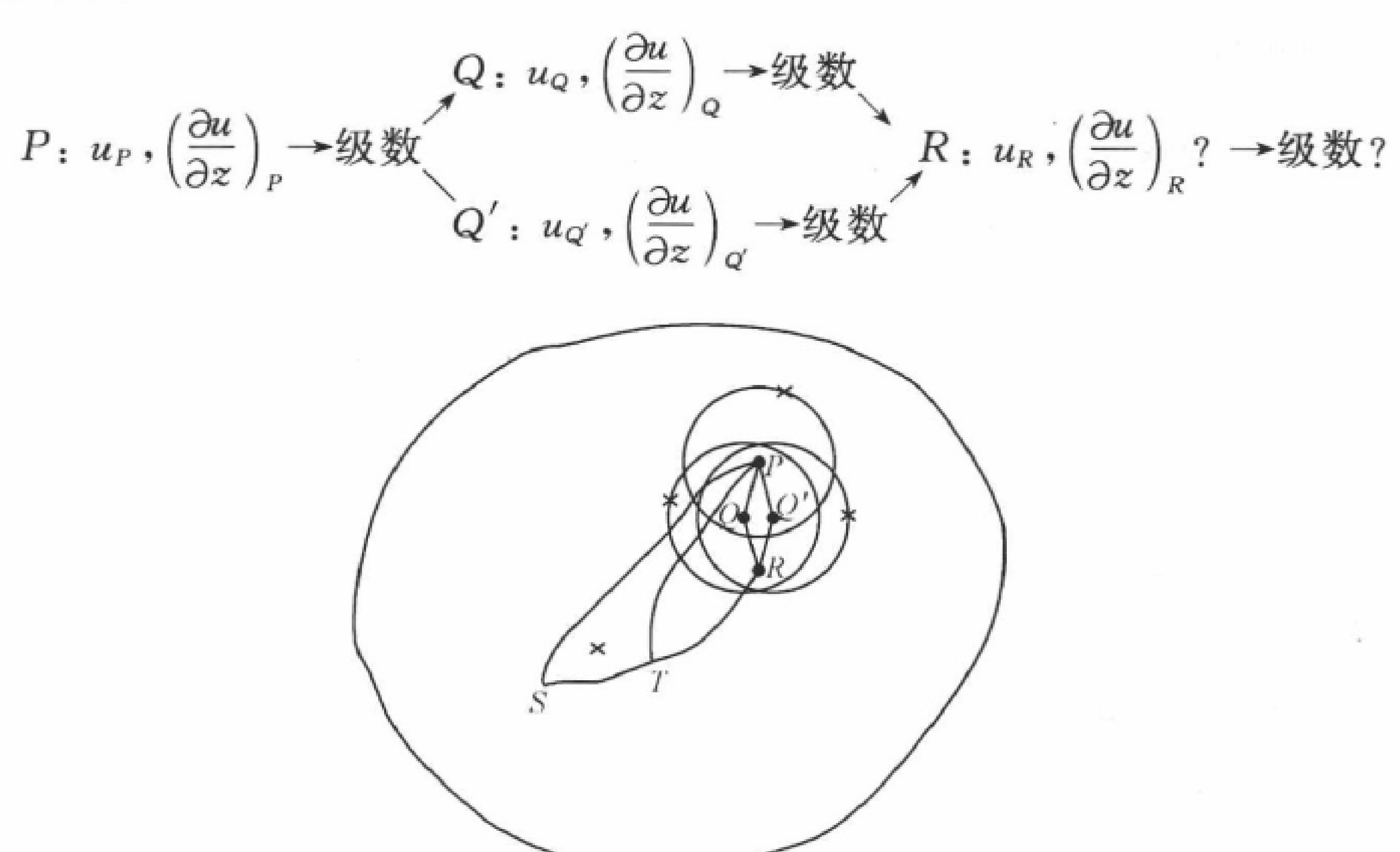
(v) 解析延拓

求得了一个常点 z_0 邻域内的级数解，可以以此作为起始幂级数，通过 Weierstrass 解析延拓方法，就可以得到整个复平面上(除奇点以外)的完全解析函数，即方程的全部常点解的表示.

12.3.4　解析延拓问题

关于解析延拓问题，还有一点需进一步加以阐述.

重要的是要知道经由两条不同路径进行延拓，是否会给出同一最后的幂级数. 例如(见下图)，



显然，由两个 Q,Q' 点算出 R 点的 $u_R,\left(\frac{\partial u}{\partial z}\right)_R$ 一定一样，即给出该点的级数也是一样的. PQR，$PQ'R$ 这样的路径称为和谐路径. 利用连续变形可以证明，凡两条路径之间不包含奇点，皆为和谐路径，如 PT 和 PRT. 相反，若其间含有奇点，则为非和谐路径，如 PS 和 PTS. 这时，沿 PS 和沿 PTS 所得 S 点之值不同；因此是多值的.

设方程在 z 点的基本解组为 u_1,u_2，通解为

$$u = c_1u_1 + c_2u_2.$$

当绕一奇点一圈回到 z 时，

$$u_1 \to u_1' = \alpha_{11}u_1 + \alpha_{12}u_2,$$
$$u_2 \to u_2' = \alpha_{21}u_1 + \alpha_{22}u_2.$$

所以，绕奇点一次相当于来一次线性变换. 如果选择适当的 v_1 和 v_2(v_1 和 v_2 是 u_1 和 u_2 的线性组合)，使得当绕奇点一次时，

$$v_1 \to v_1 e^{2\pi i\alpha_1},$$
$$v_2 \to v_2 e^{2\pi i\alpha_2};$$

α_1 和 α_2 称为**特征指数**.

12.4　正则奇点邻域内的正则解

现在来考虑方程

$$\frac{d^2 u}{dz^2} + p(z)\frac{du}{dz} + q(z)u = 0$$

在奇点邻域内的解法.

12.4.1　方程的正则奇点

对于方程的奇点 $z=z_r$，Fuchs 假定其解具有以下简单形式：

$$u = \sum_n a_n (z-z_r)^{n+\rho},$$

结果发现条件是要求 $(z-z_r)p(z)$ 和 $(z-z_r)^2 q(z)$ 在 $z=z_r$ 点是解析的. 另一方面，Frobenius 则从假定 $p(z)$ 和 $q(z)$ 在 $z=z_r$ 点满足上述条件出发，求出其解应为上述形式. 二人从不同的出发点得出同样的结果. 从而引进下面的定义.

定义　若 $(z-z_r)p(z)$ 和 $(z-z_r)^2 q(z)$ 在 z_r 点是解析的，则 $z=z_r$ 点称为方程的**正则奇点**，否则为**非正则奇点**. 同时，在 z_r 点邻域 $u=\sum_n a_n(z-z_r)^{n+\rho}$ 形式的级数解则称为方程的**正则解**.

例如，Legendre 方程（见 12.3.2 节）

$$\frac{d^2 u}{dz^2} - \frac{2z}{1-z^2}\frac{du}{dz} + \frac{3}{4(1-z^2)}u = 0,$$

有三个奇点：$z=\pm 1$ 和 $z=\infty$；它们都是正则奇点. 因为在 $z=1$ 点，

$$-(z-1)\frac{2z}{1-z^2} = \frac{2z}{1+z} = \frac{1+(z-1)}{1+\frac{1}{2}(z-1)},$$

$$(z-1)^2\frac{3}{4(1-z^2)} = \frac{3(1-z)}{4(1+z)} = \frac{3}{8}\frac{1-z}{1+\frac{1}{2}(z-1)},$$

它们都是解析的；对 $z=-1$ 点，可得类似结果. 对于 $z=\infty$ 点，作变换 $z_1=\frac{1}{z}$ 后，在 $z_1=0$ 点有

$$z_1 p_1(z_1) = z_1\left\{\frac{2z_1}{z_1^2-1}\right\} = \frac{2z_1^2}{z_1^2-1},$$

$$z_1^2 q_1(z_1) = z_1^2 \left\{ \frac{3}{4z_1^2(z_1^2-1)} \right\} = \frac{3}{4(z_1^2-1)},$$

它们也都是解析的.

12.4.2 正则解的指标方程

微分方程在正则奇点 $z=z_r$ 邻域内的**正则解**是

$$u = (z-z_r)^\rho \sum_n a_n (z-z_r)^n,$$

将上式代入一般形式的微分方程后容易得到

$$(z-z_r)^\rho \sum_n a_n (z-z_r)^{n-2} \times [(n+\rho)(n+\rho-1)$$
$$+ (n+\rho)(z-z_r)p(z) + (z-z_r)^2 q(z)] = 0.$$

由此可见,若$(z-z_r)p(z)$和$(z-z_r)^2q(z)$在 z_r 点是解析的,就可展成$(z-z_r)$的幂级数,从而可得确定系数 a_n 的递推关系.这就是前面所指出过的正则解与正则奇点之间的关系.

特别是,对于 $n=0$ 的项将会有

$$[\rho(\rho-1)+\rho p_0+q_0]a_0 = 0;$$

其中

$$p_0 = \lim_{z\to z_r}(z-z_r)p(z),$$
$$q_0 = \lim_{z\to z_r}(z-z_r)^2 q(z).$$

由于总可选择 $a_0\neq 0$(因为,若 $a_0=0$ 而 $a_1\neq 0$,则可令 $\rho'=\rho+1$,于是有 $a_0'=a_1\neq 0$),结果得到

$$\rho(\rho-1)+\rho p_0+q_0 = 0;$$

此为确定**指标** ρ 的方程,称为**指标方程**.

例 1
$$\frac{d^2u}{dz^2}-\frac{1}{z}\frac{du}{dz}+u=0,$$

$z_r=0$ 是其正则奇点;易知 $p_0=-1, q_0=0$,指标方程是

$$\rho(\rho-1)-\rho = \rho(\rho-2) = 0,$$

两个指标是 $\rho_1=2, \rho_2=0$.

例 2
$$\frac{d^2u}{dz^2}-\frac{1}{z}\frac{du}{dz}+\frac{1}{z^2}u=0,$$

对正则奇点 $z_r=0$,有指标方程

$$\rho(\rho-1)-\rho+1 = (\rho-1)^2 = 0,$$

两个指标是 $\rho_1=\rho_2=1$.

例 3
$$\frac{d^2u}{dz^2}+\frac{1+2z}{2z}\frac{du}{dz}-\frac{1}{2z^2}u=0,$$

对正则奇点 $z_r=0$，有 $p_0=\frac{1}{2}$，$q_0=-\frac{1}{2}$；指标方程是

$$\rho(\rho-1)+\frac{1}{2}\rho-\frac{1}{2}=(\rho-1)\left(\rho+\frac{1}{2}\right)=0,$$

两个指标是 $\rho_1=1$，$\rho_2=-\frac{1}{2}$.

所以，指标方程的根 ρ_1，ρ_2 有两种情形：

(i) $\rho_1-\rho_2\neq 0$ 或整数，如例 3.

在确定 a_n 的方程中分别令 $\rho=\rho_1$ 和 $\rho=\rho_2$ 可得两个解

$$u_1=(z-z_r)^{\rho_1}\sum_0^\infty (a_n)_{\rho=\rho_1}(z-z_r)^n \quad \text{和}$$

$$u_2=(z-z_r)^{\rho_2}\sum_0^\infty (a_n)_{\rho=\rho_2}(z-z_r)^n,$$

则通解为 $u=c_1u_1+c_2u_2$.

(ii) $\rho_1-\rho_2=0$（如例 2）或正整数（如例 1）.

这种情形下的两个解是：

$$u_1=(z-z_r)^{\rho_1}\sum_0^\infty (a_n)_{\rho=\rho_1}(z-z_r)^n=(u)_{\rho=\rho_1},$$

$$u_2=u_1\ln(z-z_r)+(z-z_r)^{\rho_2}\sum_{0\text{或}1}^\infty\left(\frac{\partial a_n}{\partial\rho}\right)_{\rho=\rho_2}(z-z_r)^n=\left(\frac{\partial u}{\partial\rho}\right)_{\rho=\rho_2};$$

后一求和号中，当 $\rho_2=\rho_1$ 时从 1 开始，否则从 0 开始. 顺便指出，在某些特殊情况，$\rho_1-\rho_2=$ 正整数时有可能没有对数项.

12.4.3 超几何微分方程[①][②]

下面将以**超几何微分方程**

$$z(1-z)\frac{\mathrm{d}^2u}{\mathrm{d}z^2}+\{\gamma-(\alpha+\beta+1)z\}\frac{\mathrm{d}u}{\mathrm{d}z}-\alpha\beta u=0$$

作为具体例子，进一步阐明求正则解的步骤.

(1) 三个正则奇点的 Fuchs 型方程

这个方程有三个奇点：$z_r=0,1,\infty$，它们都是正则奇点. 前两个比较显然，这里对 $z_r=\infty$ 点略作说明. 在 $z_r=\infty$，令 $z_1=\frac{1}{z}$，则得

$$p_1(z_1)=\frac{2}{z_1}-\frac{1}{z_1^2}p\left(\frac{1}{z_1}\right)=\frac{1}{z_1}\left\{\frac{(2-\gamma)z_1+(\alpha+\beta-1)}{z_1-1}\right\},$$

① Whittaker, Watson, *Modern Analysis*; Ch. XIV.

② 王竹溪、郭敦仁：《特殊函数概论》；第四章及第二章.

$$q_1(z_1)=\frac{1}{z_1^4}q\left(\frac{1}{z_1}\right)=-\frac{1}{z_1^2}\cdot\frac{\alpha\beta}{z_1-1},$$

可见此点是方程的正则奇点. 总之,凡 $p(z)$ 极点的阶 $\leqslant 1$ 和 $q(z)$ 极点的阶 $\leqslant 2$ 的,都是方程的正则奇点.

所有奇点都是正则奇点的方程称为 **Fuchs 型方程**. 由于用分式线性变换

$$z=\frac{b-c}{b-a}\frac{\zeta-a}{\zeta-c}$$

可以把任意三个点 a,b,c 依次变为 $0,1,\infty$,所以说超几何方程是具有三个正则奇点的 Fuchs 型方程的原型.

例如,球面坐标下用分离变量法解 **Helmholtz 方程**

$$(\nabla^2+k^2)V=0$$

时,得到下列方程

$$\left\{\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}}{\mathrm{d}\theta}\right)+\left[l(l+1)-\frac{m^2}{\sin^2\theta}\right]\right\}\Theta=0,$$

令 $\mu=\cos\theta, v(\mu)=\Theta(\theta)$ 后可化为**连带 Legendre 方程**:

$$(1-\mu^2)\frac{\mathrm{d}^2v}{\mathrm{d}\mu^2}-2\mu\frac{\mathrm{d}v}{\mathrm{d}\mu}+\left[l(l+1)-\frac{m^2}{1-\mu^2}\right]v=0;$$

其解是**连带 Legendre 函数**($m\neq 0$);当 $m=0$ 时,方程称为 **Legendre 方程**,其解是 **Legendre 函数**.

上述方程的三个奇点 $\mu_r=\pm 1,\infty$ 都是正则奇点. 于是,如令 $z=\frac{1}{2}(1-\mu)$,则

μ_r	1	-1	∞
$z_r=\frac{1-\mu_r}{2}$	0	1	∞

变为与超几何方程的三个正则奇点 $z_r=0,1,\infty$ 一致.

所以,下面将只讨论超几何方程的解. 此类型中其他方程的解,则可通过适当变换而推得.

(2) 指标方程及指标

对于超几何方程

$$z(1-z)\frac{\mathrm{d}^2u}{\mathrm{d}z^2}+[\gamma-(\alpha+\beta+1)z]\frac{\mathrm{d}u}{\mathrm{d}z}-\alpha\beta u=0,$$

$$p(z)=\frac{\gamma-(\alpha+\beta+1)z}{z(1-z)}=\frac{\gamma}{z}+\frac{1+\alpha+\beta-\gamma}{z-1},$$

$$q(z)=-\frac{\alpha\beta}{z(1-z)}=-\frac{\alpha\beta}{z}+\frac{\alpha\beta}{z-1};$$

在 $z_r=0$ 和 $z_r=1$ 的指标方程及指标可分别求得为

$$z_r=0:\ \rho(\rho-1)+\rho\gamma=\rho[\rho-(1-\gamma)]=0,$$
$$\rho_1=0 \quad 和 \quad \rho_2=1-\gamma;$$
$$z_r=1:\ \rho(\rho-1)+\rho(1+\alpha+\beta-\gamma)=\rho[\rho-(\gamma-\alpha-\beta)]=0,$$
$$\rho_1=0 \quad 和 \quad \rho_2=\gamma-\alpha-\beta.$$

在 $z_r=\infty$，令 $z_1=\dfrac{1}{z}$后，超几何方程变为

$$z_1(1-z_1)\frac{\mathrm{d}^2u}{\mathrm{d}z_1^2}+[(1-\alpha-\beta)-(2-\gamma)z_1]\frac{\mathrm{d}u}{\mathrm{d}z_1}+\frac{\alpha\beta}{z_1}=0,$$

$$\lim_{z_1\to 0}[z_1p_1(z_1)]=\lim_{z_1\to 0}\frac{(1-\alpha-\beta)-(2-\gamma)z_1}{1-z_1}=1-\alpha-\beta,$$

$$\lim_{z_1\to 0}[z_1^2q_1(z_1)]=\lim_{z_1\to 0}\frac{\alpha\beta}{1-z_1}=\alpha\beta;$$

可求得指标方程为

$$\rho(\rho-1)+\rho(1-\alpha-\beta)+\alpha\beta=(\rho-\alpha)(\rho-\beta)=0,$$

两指标为 α 和 β.

前面曾指出，正则奇点邻域内的正规解可由该奇点的两个指标给出. 因此，超几何方程的全部解可用 P 符号表示为

$$u(z)=\mathrm{P}\begin{Bmatrix}0 & 1 & \infty & \\ 0 & 0 & \alpha; & z \\ 1-\gamma & \gamma-\alpha-\beta & \beta & \end{Bmatrix}.$$

(3) $z_r=0$ 邻域内的正则解. 超几何函数

现在较详细地来求超几何方程在 $z_r=0$ 点的正则解. 令(注意 $a_0\neq 0$)

$$u=(z-0)^\rho\{a_0+a_1(z-0)+a_2(z-0)^2+\cdots\}=\sum_{n=0}^{\infty}a_nz^{\rho+n},$$

将此解代入微分方程

$$\overset{-1}{z}(1-\overset{0}{z})\frac{\mathrm{d}^2u}{\mathrm{d}z^2}+\{\overset{-1}{\gamma}-(\alpha+\beta+1)\overset{0}{z}\}\frac{\mathrm{d}u}{\mathrm{d}z}-\overset{0}{\alpha}\beta u=0 \quad ^{2阶步},$$

注意到它共有二阶步，将其标清楚后，即刻可写出

$$\sum_{n=0}^{\infty}a_n\{(\rho+n)(\rho+n-1)+\gamma(\rho+n)\}z^{\rho+n-1}$$
$$-\sum_{n=0}^{\infty}a_n\{(\rho+n)(\rho+n-1)+(\alpha+\beta+1)(\rho+n)+\alpha\beta\}z^{\rho+n}=0.$$

或者，在第一行中将 $n=0$ 项单独写出，再将通项中的 n 换为 $n+1$，并经简单代数运算后上式变成

$$a_0\{\rho(\rho-1)+\gamma\rho\}z^{\rho-1}+\sum_{n=0}^{\infty}a_{n+1}(\rho+n+1)(\rho+n+\gamma)z^{\rho+n}$$

$$-\sum_{n=0}^{\infty}a_n(\rho+n+\alpha)(\rho+n+\beta)z^{\rho+n}=0.$$

自变量同次幂的系数和应为零. 于是,$z^{\rho-1}$ 的系数为零(及 $a_0\neq0$)给出**指标方程**

$$\rho(\rho-1)+\gamma\rho=0,$$

可解出两个指标 $\rho_1=0$ 和 $\rho_2=1-\gamma$,正如前面曾求得的. 通项 $z^{\rho+n}$ 的系数为零则给出确定 a_n 的递推关系

$$a_{n+1}=\frac{(\rho+n+\alpha)(\rho+n+\beta)}{(\rho+n+1)(\rho+n+\gamma)}a_n,\quad n=0,1,2,\cdots.$$

因此,给定 $a_0\neq0$,即可求得 a_1,由 a_1 得 a_2,…,现在将系数 a_n 具体写出为

$$a_1=\frac{(\rho+\alpha)(\rho+\beta)}{(\rho+1)(\rho+\gamma)}a_0,$$

$$a_2=\frac{(\rho+\alpha)(\rho+\alpha+1)(\rho+\beta)(\rho+\beta+1)}{(\rho+1)(\rho+2)(\rho+\gamma)(\rho+\gamma+1)}a_0,$$

$$\vdots$$

$$a_n=\frac{(\rho+\alpha)_n(\rho+\beta)_n}{(\rho+1)_n(\rho+\gamma)_n}a_0;$$

其中

$$(\lambda)_0=1,\quad(\lambda)_n=\lambda(\lambda+1)\cdots(\lambda+n-1)\quad(n\geqslant1).$$

于是,此正则解可写成(令 $a_0=1$)

$$u=z^{\rho}\sum_{n=0}^{\infty}a_nz^n=z^{\rho}\sum_{n=0}^{\infty}\frac{(\rho+\alpha)_n(\rho+\beta)_n}{(\rho+1)_n(\rho+\gamma)_n}z^n.$$

所以,超几何方程在 $z_r=0$ 点邻域内有两个正则解,分别对应于指标 $\rho_1=0$ 和 $\rho_2=1-\gamma$.

对于 $\rho_1=0$ 的**第一解**为

$$u_1=\left\{1+\frac{\alpha\cdot\beta}{1\cdot\gamma}z+\frac{\alpha(\alpha+1)\cdot\beta(\beta+1)}{1\cdot2\cdot\gamma(\gamma+1)}z^2+\cdots\right\}$$

$$=\sum_{n=0}^{\infty}\frac{(\alpha)_n(\beta)_n}{n!(\gamma)_n}z^n\equiv\mathrm{F}(\alpha,\beta;\gamma;z);$$

其中用了 $(1)_n=1\cdot2\cdot3\cdot\cdots\cdot n=n!$ (n 的**阶乘**). 这个级数称为**超几何级数**,其收敛半径为 1,在收敛圆 $|z|<1$ 内代表一个解析函数;它可解析延拓到(可能除 $z=1$ 和 $z=\infty$ 两点之外的)全 z 平面而称为**超几何函数**.

应注意到,对超几何函数有

$$\mathrm{F}(\alpha,\beta;\gamma;0)=1,$$

$$\mathrm{F}(\alpha,\beta;\gamma;z)=\mathrm{F}(\beta,\alpha;\gamma;z).$$

对于 $\rho_2=1-\gamma$ 的**第二解**，当 γ 不等于整数时为

$$u_2 = z^{\rho_2}\sum_{n=0}^{\infty}\frac{(\rho_2+\alpha)_n(\rho_2+\beta)_n}{(\rho_2+\gamma)_n(\rho_2+1)_n}z^n$$

$$= z^{1-\gamma}\sum_{n=0}^{\infty}\frac{(1-\gamma+\alpha)_n(1-\gamma+\beta)_n}{(1)_n(2-\gamma)_n}z^n$$

$$= z^{1-\gamma}\mathrm{F}(1-\gamma+\alpha,1-\gamma+\beta;2-\gamma;z).$$

还应指出，对于上述两个正则解 u_1 和 u_2，当 $\gamma>1$ 时($1-\gamma\neq 0$ 或整数)，在 $z=0$ 点的 u_1 有限而 u_2 无限. 一个解良态而另一个解病态，这是正则解的一般现象.

有了两个线性无关解 u_1 和 u_2，于是，方程的通解可表达为

$$u = c_1u_1 + c_2u_2.$$

超几何方程在 $z_r=0$ 邻域内的两个正则解

$$u_1 = \mathrm{F}(\alpha,\beta;\gamma;z),$$

$$u_2 = z^{1-\gamma}\mathrm{F}(1-\gamma+\alpha,1-\gamma+\beta;2-\gamma;z);$$

其中 u_2 是在 γ 不等于整数的假定下得出的. 若 $\gamma=1$，则显然有 $u_2=u_1$，因为这时两个指标 $\rho_2=1-\gamma$ 与 $\rho_1=0$ 之值相同. 对于此情况，可采用先从 u_2 减去 u_1，然后再取 $\gamma\to 1$(即 $\rho\to 0$)的极限过程，这样可获得对数项；也就是说，相当于 $\lim\limits_{\rho\to 0}\left(\dfrac{\partial u_2}{\partial\rho}\right)$. 结果得到 $\gamma=1$(即 $\rho_2=\rho_1=0$)时两个正则解应是

$$u_1 = \mathrm{F}(\alpha,\beta;1;z),$$

$$u_2 = (\ln z)\mathrm{F}(\alpha,\beta;1;z) + \left\{\frac{\partial}{\partial\rho}\mathrm{F}(\rho+\alpha,\rho+\beta;\rho+\gamma;z)\right\}\bigg|_{\rho=0};$$

后者还可写成

$$u_2 = (\ln z)u_1 + \sum_{n=1}^{\infty}\left(\frac{\partial a_n}{\partial\rho}\right)_{\rho=0}z^n,$$

$$a_n = \frac{(\rho+\alpha)_n(\rho+\beta)_n}{n!(\rho+\gamma)_n},$$

这正是前面曾经指出过的.

(4) $z_r=1$ 邻域内的正则解

至于 $z_r=1$ 邻域内的解，可通过 $z\to 1-z$ 的变换来求. 经过变换后，所得微分方程仍是超几何方程，只是参数有所不同

$$z\to 1-z:\ \alpha,\beta,\gamma\to\alpha,\beta,1+\alpha+\beta-\gamma;$$

新方程的奇点 0 对应于原方程的奇点 1，新方程的两个指标则为 0 和 $\gamma-\alpha-\beta$，因而可由上述的 u_1，u_2 立即得出 $z_r=1$ 邻域内对应的两个正则解为

$$u_3 = \mathrm{F}(\alpha,\beta;1+\alpha+\beta-\gamma;1-z),$$

$$u_4 = (1-z)^{\gamma-\alpha-\beta}\mathrm{F}(\gamma-\alpha,\gamma-\beta;1-\alpha-\beta+\gamma;1-z);$$

这里，对 u_4 假定 $\gamma-\alpha-\beta$ 不等于整数.

(5) $z_r=\infty$ 邻域内的正则解

对于 $z_r=\infty$ 邻域内的情况，首先通过 $z=\dfrac{1}{z_1}$ 的变换得到

$$\frac{d^2u}{dz_1^2}+p_1(z_1)\frac{du}{dz_1}+q_1(z_1)u=0,$$

$$p_1(z_1)=\frac{(1-\alpha-\beta)-(2-\gamma)z_1}{z_1(1-z_1)},$$

$$q_1(z_1)=\frac{\alpha\beta}{z_1^2(1-z_1)};$$

两指标为 α,β. 再通过令 $u=z_1^\alpha v$ 的变换，容易得到

$$\frac{d^2v}{dz_1}+p_v(z_1)\frac{dv}{dz_1}+q_v(z_1)v=0,$$

$$p_v(z_1)=p_1(z_1)+\frac{2\alpha}{z_1}=\frac{(1+\alpha-\beta)-(2+2\alpha-\gamma)z_1}{z_1(1-z_1)},$$

$$q_v(z_1)=q_1(z_1)+\frac{\alpha}{z_1}p_1(z_1)+\frac{\alpha(\alpha-1)}{z_1^2}=-\frac{\alpha(1+\alpha-\gamma)}{z_1(1-z_1)};$$

此方程亦为超几何方程，两指标为 $0,\beta-\alpha$；相应参数是：$\alpha,\alpha-\gamma+1,\alpha-\beta+1$. 因而可由上述的 u_1,u_2 立即得出 $z_r=\infty$ 邻域内对应的两个正则解为

$$u_5=(-z)^{-\alpha}\mathrm{F}(\alpha,\alpha-\gamma+1;\alpha-\beta+1;z^{-1}),$$
$$u_6=(-z)^{-\beta}\mathrm{F}(\beta,\beta-\gamma+1;\beta-\alpha+1;z^{-1});$$

这里假定 $\beta-\alpha$ 不等于整数，而 z 前的负号因子是为了以后讨论时的方便引进的.

(6) 边值问题

在三个正则奇点邻域各有两个正则解，六个解中只有两个线性无关解；这些解通过解析延拓而相互联系着.

另一方面，若选定线性无关解 u_1,u_2，则方程的通解可表达为

$$u=c_1u_1+c_2u_2,$$

对于具体问题的解，则由初值 $(u)_A,\left(\dfrac{du}{dz}\right)_A$ 或由边值 $(u)_A,(u)_B$ 来确定解中常数. 以上是指常点(A 或 A 和 B)而言.

由于在奇点处给定的一个边界条件，它要比常点处的一个边界条件对解的限制多一些内涵. 因此，若在奇点处给定两个边界条件，则方程可能无解；假定有解的话，则微分方程中的参数必须是某些特殊值，或者参数间必须有某种特殊关系，此即所谓**本征值问题**.

例如，若在奇点 S 处给定一个弱的条件“u 是有限的”，即可求出通解 $u=c_1u_1+c_2u_2$ 中的一个常数. 由于一般情况是：一个解如 u_1 在 S 点有限而另一个解如 u_2

在 S 点无限，所以条件要求 $c_2=0$.

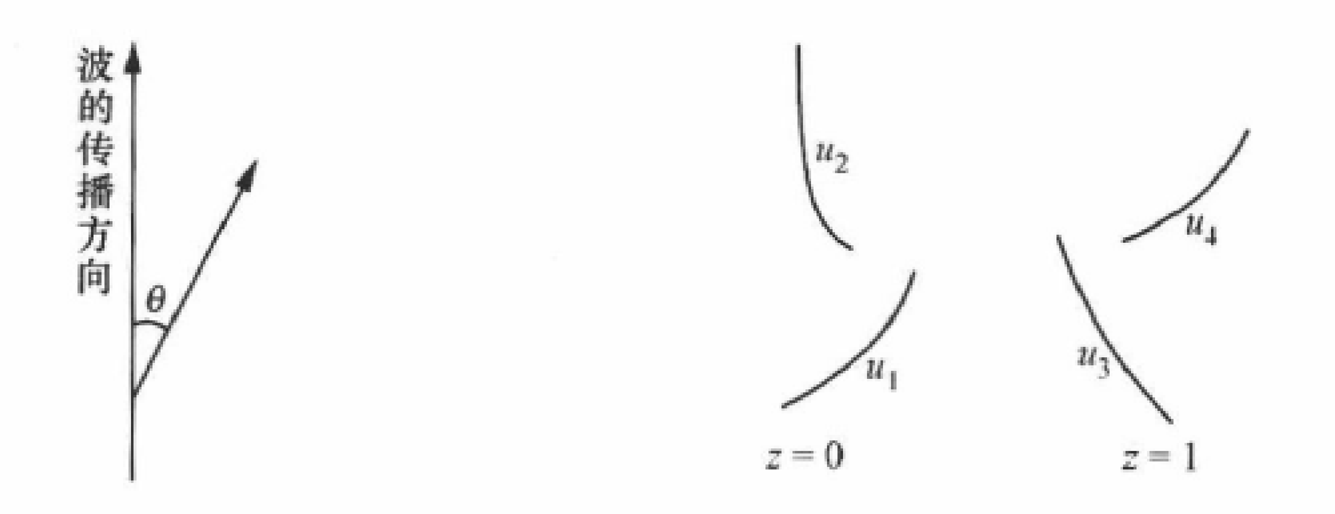
下面以连带 Legendre 方程

$$(1-\zeta^2)\frac{\mathrm{d}^2u}{\mathrm{d}\zeta^2}-2\zeta\frac{\mathrm{d}u}{\mathrm{d}\zeta}+\left\{l(l+1)-\frac{m^2}{1-\zeta^2}\right\}u=0$$

$(\zeta=\cos\theta)$为例来稍作说明. 通过令 $z=\frac{1}{2}(1-\zeta)$，可使之化为超几何方程.

由(下左图)$z=\frac{1}{2}(1-\cos\theta)$，$z=0,\theta=0,z=1,\theta=\pi$. 从物理现象知道，$u$ 在 $z=0,1$ 二点一定得有限. 由于 $u=c_1u_1+c_2u_2$：(见下中图及本小节之(3)) u_1 在 $z=0$ 有限，u_2 在 $z=0$ 无限. 所以，u 在 $z=0$ 点有限导致 $c_2=0$. 所以，

$$u=c_1\mathrm{F}(\alpha,\beta;\gamma;z).$$



另外，u 在 $z=1$ 点有限导致(见上右图及本小节之(4)中的 u_4)

$$\mathrm{Re}(\gamma-\alpha-\beta)>0,$$

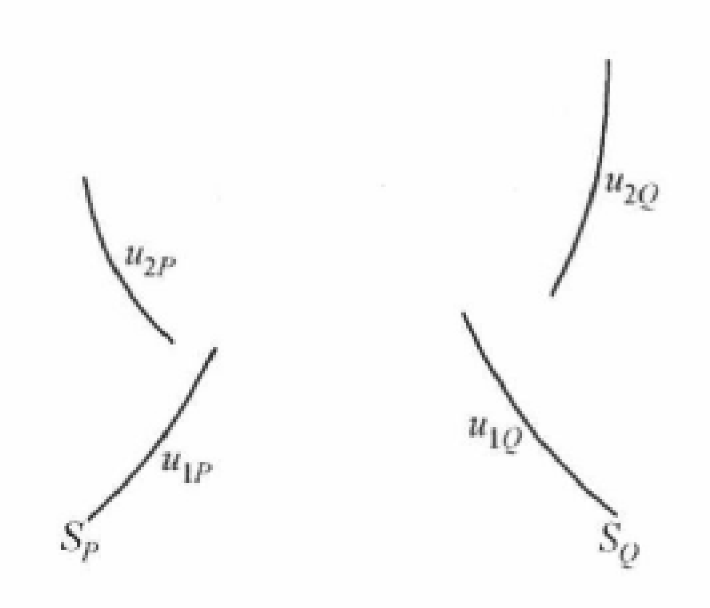
$\mathrm{Re}\lambda$ 代表 λ 的实数部分；这样得到方程的参数之间的一个关系，即本征值问题.

总之，凡是有超定边界条件，就可确定微分方程中的参数或参数间的关系. 如何确定的问题即所谓本征值问题.

(7) 解析延拓问题

下面来讨论解析延拓问题.

前面曾经指出，正则奇点邻域内的两个正则解，一般情况下，其中一个解在奇点处是良态的，另一个解则是病态的，如下图所示.



若 S_P 和 S_Q 是二正则奇点：

在 S_P 的二解是 u_{1P} 及 u_{2P}；

在 S_Q 的二解是 u_{1Q} 及 u_{2Q}.

则有

$$u = c_{1P}u_{1P} + c_{2P}u_{2P} = c_{1Q}u_{1Q} + c_{2Q}u_{2Q}.$$

所谓解析延拓问题，就是知道 u_{1P} 及 u_{2P}，如何去求它们在 S_Q 邻域的表达式的问题；即求

$$u_{1P} = c_{1Q}^{1P}u_{1Q} + c_{2Q}^{1P}u_{2Q},$$

也即确定 c_{1Q}^{1P} 及 c_{2Q}^{1P} 的问题.

当 u 在 S_P 及 S_Q 是有限时，则由

$$u = c_{1P}u_{1P} = c_{1P}(c_{1Q}^{1P}u_{1Q} + c_{2Q}^{1P}u_{2Q})$$

可知，要求 u_{2Q} 为有限导致方程参数需满足的条件. 所以，本征值问题可看作解析延拓问题的特例.

解析延拓问题的重要性在于：若已知 u_{1P} 及 u_{2P}，并知 $u=c_{1p}u_{1P}+c_{2p}u_{2P}$，欲求离 S_Q 很近一点的 u 值时，如果知道 u_{1Q} 及 u_{2Q}，并知 u_{1P} 和 u_{2P} 在 S_Q 邻域的延拓表达式，会得到 u 在该邻域收敛很快的级数，很容易计算；否则的话，直接用 u_{1P} 和 u_{2P} 会收敛很慢，不便计算.

关于这类解析延拓问题，一般采用(i)级数变换的方法，或(ii)周线积分的方法.

超几何函数的延拓问题已经做出，结果举例如下[①][②]：

在 $z_r=\infty$ 邻域内的两个正则解是

$$u_5 = (-z)^{-\alpha}\mathrm{F}\left(\alpha,\alpha-\gamma+1;\alpha-\beta+1;\frac{1}{z}\right),$$

$$u_6 = (-z)^{-\beta}\mathrm{F}\left(\beta,\beta-\gamma+1;\beta-\alpha+1;\frac{1}{z}\right).$$

$u_1=\mathrm{F}(\alpha,\beta;\gamma;z)$ 在 $z_r=\infty$ 邻域内的解析延拓是

$$u_1 = \mathrm{F}(\alpha,\beta;\gamma;z) = c_{1,5}u_5 + c_{1,6}u_6,$$

$$c_{1,5} = \frac{\Gamma(\gamma)\Gamma(\beta-\alpha)}{\Gamma(\gamma-\alpha)\Gamma(\beta)},\quad c_{1,6} = \frac{\Gamma(\gamma)\Gamma(\alpha-\beta)}{\Gamma(\gamma-\beta)\Gamma(\alpha)};$$

其中 $\Gamma(\zeta)$ 是 12.2.6 节中所定义的 Γ 函数.

在 $z_r=1$ 邻域内的两个正则解是

$$u_3 = \mathrm{F}(\alpha,\beta;\alpha+\beta-\gamma+1;1-z),$$

$$u_4 = (1-z)^{\gamma-\alpha-\beta}\mathrm{F}(\gamma-\alpha,\gamma-\beta;\gamma-\alpha-\beta+1;1-z).$$

---

① Whittaker, Watson, *Modern Analysis*, pp. 289,291.

② 王竹溪，郭敦仁：《特殊函数概论》，4.8 节.

$u_1 = \mathrm{F}(\alpha,\beta;\gamma;z)$在 $z_r=1$ 邻域内的解析延拓是

$$u_1 = \mathrm{F}(\alpha,\beta;\gamma;z) = c_{1,3}u_3 + c_{1,4}u_4,$$

$$c_{1,3} = \frac{\Gamma(\gamma-\alpha-\beta)\Gamma(\gamma)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)},\quad c_{1,4} = \frac{\Gamma(\alpha+\beta-\gamma)\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\beta)}.$$

12.4.4 Legendre 方程[①][②]

(1) 连带 Legendre 方程与 Legendre 方程

前面[12.4.3 节之(1)]曾经引进**连带 Legendre 方程**

$$(1-\zeta^2)\frac{\mathrm{d}^2 v}{\mathrm{d}\zeta^2} - 2\zeta\frac{\mathrm{d}v}{\mathrm{d}\zeta} + \left[l(l+1) - \frac{m^2}{1-\zeta^2}\right]v = 0,$$

它的三个正则奇点是 $\zeta_r = \pm 1, \infty$，将它写成

$$\frac{\mathrm{d}^2 v}{\mathrm{d}\zeta^2} + \left\{\frac{1}{\zeta+1} + \frac{1}{\zeta-1}\right\}\frac{\mathrm{d}v}{\mathrm{d}\zeta} + \left\{\frac{m^2/2}{\zeta+1} + \frac{-m^2/2}{\zeta-1} - l(l+1)\right\}\frac{v}{(\zeta+1)(\zeta-1)} = 0,$$

容易求得 $p_0(\pm 1) = 1, q_0(\pm 1) = -\frac{m^2}{4}$以及

$$p_0(\infty) = \lim_{\zeta\to\infty}\{2 - \zeta p(\zeta)\} = 0,$$

$$q_0(\infty) = \lim_{\zeta\to\infty}\{\zeta^2 q(\zeta)\} = -l(l+1);$$

于是，相应指标由指标方程 $\rho(\rho-1) + p_0\rho + q_0 = 0$ 求得分别为$\pm\frac{m}{2}$以及 $l+1, -l$.

连带 Legendre 方程的全部解可用 P 符号表达为

$$v(\zeta) = \mathrm{P}\left\{\begin{matrix} +1 & -1 & \infty & \\ \frac{m}{2} & \frac{m}{2} & l+1; & \zeta \\ -\frac{m}{2} & -\frac{m}{2} & -l & \end{matrix}\right\}.$$

应该指出，当 $m=0$ 时，连带 Legendre 方程化为 **Legendre 方程**.

(2) P 方程的变换[③]

利用 **P 方程的变换**，可将连带 Legendre 方程的解用超几何方程的解予以表达.

首先，通过自变量的线性变换 $z = \frac{1}{2}(1-\zeta)$，可将 $\zeta_r = +1, -1, \infty$变为 $z_r = 0, 1, \infty$，同时相应指标不变.（一般而言，分式线性变换只改变奇点而保持相应指标不

① Whittaker, Watson, *Modern Analysis*, Ch. XV, XIV.

② 王竹溪，郭敦仁：《特殊函数概论》；第五章及第二章.

③ 王竹溪，郭敦仁：《特殊函数概论》，2.9 节.

变，）因此有

$$v(\zeta)=\mathrm{P}\begin{Bmatrix}1 & -1 & \infty & \\ \dfrac{m}{2} & \dfrac{m}{2} & l+1; & \zeta \\ -\dfrac{m}{2} & -\dfrac{m}{2} & -l & \end{Bmatrix}=\mathrm{P}\begin{Bmatrix}0 & 1 & \infty & \\ \dfrac{m}{2} & \dfrac{m}{2} & l+1; & \dfrac{1-\zeta}{2} \\ -\dfrac{m}{2} & -\dfrac{m}{2} & -l & \end{Bmatrix}.$$

另一方面，关于指标的变换，一般有

$$\mathrm{P}\begin{Bmatrix}a & b & \infty & \\ \alpha_1+k_a & \beta_1+k_b & \gamma_1-k_a-k_b; & z \\ \alpha_2+k_a & \beta_2+k_b & \gamma_2-k_a-k_b & \end{Bmatrix}$$

$$=(z-a)^{k_a}(z-b)^{k_b}\mathrm{P}\begin{Bmatrix}a & b & \infty & \\ \alpha_1 & \beta_1 & \gamma_1; & z \\ \alpha_2 & \beta_2 & \gamma_2 & \end{Bmatrix}.$$

对于现在的情况，最后得到

$$v(\zeta)=\mathrm{P}\begin{Bmatrix}1 & -1 & \infty & \\ \dfrac{m}{2} & \dfrac{m}{2} & l+1; & \zeta \\ -\dfrac{m}{2} & -\dfrac{m}{2} & -l & \end{Bmatrix}=\mathrm{P}\begin{Bmatrix}0 & 1 & \infty & \\ \dfrac{m}{2} & \dfrac{m}{2} & l+1; & \dfrac{1-\zeta}{2} \\ -\dfrac{m}{2} & -\dfrac{m}{2} & -l & \end{Bmatrix}$$

$$=2^{-m}(1-\zeta^2)^{m/2}\mathrm{P}\begin{Bmatrix}0 & 1 & \infty & \\ 0 & 0 & l+m+1; & \dfrac{1-\zeta}{2} \\ -m & -m & -l+m & \end{Bmatrix};$$

其中用了 $\left(\dfrac{1-\zeta}{2}\right)^{m/2}\left(1-\dfrac{1-\zeta}{2}\right)^{m/2}=2^{-m}(1-\zeta^2)^{m/2}$.

与超几何方程的 P 方程

$$u(z)=\mathrm{P}\begin{Bmatrix}0 & 1 & \infty & \\ 0 & 0 & \alpha; & z \\ 1-\gamma & \gamma-\alpha-\beta & \beta & \end{Bmatrix}$$

比较，得出 $\alpha=l+m+1,\beta=-l+m,\gamma=1+m,z=\dfrac{1-\zeta}{2}$.

(3) 连带 Legendre 函数与 Legendre 函数

按照上一小节的结果，得到连带 Legendre 方程的一个解是**连带 Legendre 函数**

$$\mathrm{P}_l^m(\zeta)=(-)^m\frac{(1-\zeta^2)^{m/2}}{2^m m!}\frac{\Gamma(l+m+1)}{\Gamma(l-m+1)}\mathrm{F}\left(l+m+1,-l+m;1+m;\frac{1-\zeta}{2}\right),$$

而 Legendre 方程的相应解是 **Legendre 函数**

$$\mathrm{P}_l(\zeta) = \mathrm{F}\left(l+1, -l; 1; \frac{1-\zeta}{2}\right).$$

注意到超几何函数

$$\mathrm{F}(\alpha,\beta;\gamma;z) = \sum_{n=0}^{\infty} \frac{(\alpha)_n(\beta)_n}{n!(\gamma)_n} z^n$$

之间的下列关系：

$$\frac{\mathrm{d}}{\mathrm{d}z}\mathrm{F}(\alpha,\beta;\gamma,z) = \sum_{n=1}^{\infty} \frac{(\alpha)_n(\beta)_n}{(n-1)!(\gamma)_n} z^{n-1} = \frac{\alpha\beta}{\gamma}\sum_{n=0}^{\infty} \frac{(\alpha+1)_n(\beta+1)_n}{n!(\gamma+1)_n} z^n$$

$$= \frac{\alpha\beta}{\gamma}\mathrm{F}[\alpha+1,\beta+1;\gamma+1;z],$$

$$\frac{\mathrm{d}^m}{\mathrm{d}z^m}\mathrm{F}(\alpha,\beta;\gamma;z) = \frac{(\alpha)_m(\beta)_m}{(\gamma)_m}\mathrm{F}(\alpha+m,\beta+m;\gamma+m;z);$$

容易得到①(当 m 为正整数时)

$$\mathrm{P}_l^m(\zeta) = (-)^m(1-\zeta^2)^{\frac{m}{2}} \frac{\mathrm{d}^m}{\mathrm{d}\zeta^m}\mathrm{P}_l(\zeta).$$

这也就部分说明了 $\mathrm{P}_l^m(\zeta)$ 中引进 $(2^m m!)^{-1}\dfrac{\Gamma(l+m+1)}{\Gamma(l-m+1)}$ 系数的原因.

(4) Legendre 多项式②

前面(12.3.2 节)曾经求过 Legendre 方程在其常点 $\zeta=0$ 点邻域内的级数解 $v(\zeta) = \sum a_n\zeta^n$，得到下列递推关系

$$a_{n+2} = \frac{(n-l)(n+l+1)}{(n+1)(n+2)}a_n,$$

容易求得

$$v(\zeta) = a_0\sum \frac{\left(-\frac{l}{2}\right)_n\left(\frac{l+1}{2}\right)_n}{n!\left(\frac{1}{2}\right)_n}\zeta^{2n} + a_1\zeta\sum \frac{\left(\frac{1-l}{2}\right)_n\left(\frac{l}{2}+1\right)_n}{n!\left(\frac{3}{2}\right)_n}\zeta^{2n}$$

$$= a_0\mathrm{F}\left(-\frac{l}{2},\frac{l+1}{2};\frac{1}{2};\zeta^2\right) + a_1\zeta\mathrm{F}\left(\frac{1-l}{2},\frac{l}{2}+1;\frac{3}{2};\zeta^2\right),$$

其中 $\mathrm{F}(\alpha,\beta;\gamma;z)$ 是超几何函数.

显而易见，当 $l=0,1,2,\cdots$ 时，有 $a_{l+2}=a_{l+4}=\cdots=0$. 故若 l 为偶数，则前者(含 a_0 者)为 l 次多项式，而后者(含 a_1 者)为无穷级数；若 l 为奇数，则后者为 l 次多项

---

① 这里应用了 $\Gamma(l\pm m+1)=(l\pm m)!$ 和

$$(-l)_m = (-l)(-l+1)\cdots(-l+m-1) = (-)^m l(l-1)\cdots(l-m+1) = (-)^m l!/(l-m)!,$$

$$(l+1)_m = (l+m)!/l!, \quad (1)_m = m!;$$

以及由 $z=\dfrac{1-\zeta}{2}$ 得 $\dfrac{\mathrm{d}^m}{\mathrm{d}z^m} = (-)^m 2^{-m}\dfrac{\mathrm{d}^m}{\mathrm{d}\zeta^m}$.

② 王竹溪，郭敦仁：《特殊函数概论》，5.2 节，5.3 节.

式，而前者为无穷级数. 由前述 Legendre 函数（l 可以是非整数）可知，当 $l=0,1,2,\cdots$ 时，它变为 **Legendre 多项式**，其最高次项 ζ^l 的系数是

$$\frac{(l+1)_l(-l)_l}{l!(1)_l}\left(-\frac{1}{2}\right)^l=\frac{(2l)!}{2^l(l!)^2}.$$

于是由上述递推关系（倒推）可得

$$\begin{aligned}\mathrm{P}_l(\zeta)&=\frac{(2l)!}{2^l(l!)^2}\zeta^l\left\{1-\frac{l(l-1)}{2(2l-1)}\zeta^{-2}+\frac{l(l-1)(l-2)(l-3)}{2\cdot4\cdot(2l-1)(2l-3)}\zeta^{-4}-\cdots\right\}\\&=\frac{(2l)!}{2^l(l!)^2}\zeta^l\mathrm{F}\left(-\frac{l}{2},\frac{1-l}{2};\frac{1}{2}-l;\zeta^{-2}\right)\\&=\frac{1}{2^l}\sum_{r=0}^{\left[\frac{l}{2}\right]}(-)^r\frac{(2l-2r)!}{r!(l-r)!(l-2r)!}\zeta^{l-2r}.\end{aligned}$$

其中 $\left[\frac{l}{2}\right]$ 表示 $\leqslant\frac{l}{2}$ 的最大整数.

Legendre 多项式还可表示成

$$\mathrm{P}_l(\zeta)=\frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}\zeta^l}(\zeta^2-1)^l,$$

称为 **Rodrigues 公式**. 通过二项式展开后直接求导可得上面的最后一个展开式，此即证明.

下面列出前几个 Legendre 多项式：

$\mathrm{P}_0(x)=1$，

$\mathrm{P}_1(x)=x$，

$\mathrm{P}_2(x)=\frac{1}{2}(3x^2-1)$，

$\mathrm{P}_3(x)=\frac{1}{2}(5x^3-3x)$，

$\mathrm{P}_4(x)=\frac{1}{8}(35x^4-30x^2+3)$，

$\mathrm{P}_5(x)=\frac{1}{8}(63x^5-70x^3+15x)$，

$\mathrm{P}_6(x)=\frac{1}{16}(231x^6-315x^4+105x^2-5)$，

$\mathrm{P}_7(x)=\frac{1}{16}(429x^7-693x^5+315x^3-35x)$.

(5) 递推关系①②

首先，列出 $P_l(x)$ 的主要**递推关系**如下：

$$(l+1)P_{l+1}+lP_{l-1}=(2l+1)xP_l,$$

$$\frac{d}{dx}P_{l+1}=x\frac{d}{dx}P_l+(l+1)P_l,$$

$$\frac{d}{dx}P_{l-1}=x\frac{d}{dx}P_l-lP_l,$$

$$\frac{d}{dx}P_{l+1}-\frac{d}{dx}P_{l-1}=(2l+1)P_l,$$

$$(1-x^2)\frac{d}{dx}P_l=lP_{l-1}-lxP_l,$$

$$(1-x^2)\frac{d}{dx}P_l=(l+1)xP_l-(l+1)P_{l+1}.$$

利用 $P_l(x)$ 的**生成函数**

$$(1-2xt+t^2)^{-\frac{1}{2}}=\sum_{l=0}^{\infty}P_l(x)t^l,$$

通过两边对 t 或对 x 求导数和一些代数运算，并比较两边 t 的同次方的系数，以及其他简单运算，即可得以上诸关系；具体证明从略.

其次，列出 $P_l^m(x)$ 的主要递推关系如下：

$$(2l+1)xP_l^m=(l+m)P_{l-1}^m+(l-m+1)P_{l+1}^m,$$

$$P_{l-1}^m=xP_l^m+(l-m+1)(1-x^2)^{\frac{1}{2}}P_l^{m-1},$$

$$P_{l+1}^m=xP_l^m-(l+m)(1-x^2)^{\frac{1}{2}}P_l^{m-1},$$

$$(1-x^2)^{\frac{1}{2}}P_l^{m+1}=(l-m+1)P_{l+1}^m-(l+m+1)xP_l^m,$$

$$\begin{aligned}(1-x^2)\frac{d}{dx}P_l^m(x)&=(l+1)xP_l^m(x)-(l-m+1)P_{l+1}^m(x)\\&=(l+m)P_{l-1}^m(x)-lxP_l^m(x).\end{aligned}$$

利用下列公式

$$P_l^m(x)=(-)^m(1-x^2)^{\frac{m}{2}}\frac{d^m}{dx^m}P_l(x),$$

可以很容易地从 $P_l(x)$ 的递推关系求出 $P_l^m(x)$ 的上述递推关系；具体推导从略.

---

① 王竹溪，郭敦仁：《特殊函数概论》，5.5 节，5.13 节.

② H. Bateman, *Partial differential Equations of Mathematical Physics*, Cambridge, New York, 1994, p. 360.

(6) 正交归一关系[1][2]

$\mathrm{P}_l^m(x)$满足下列**正交归一关系**

$$\int_{-1}^{1}\mathrm{P}_l^m\mathrm{P}_{l'}^m\mathrm{d}x=\frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\delta_{ll'},$$

$$\int_{-1}^{1}\mathrm{P}_l^m\mathrm{P}_l^{m'}\frac{\mathrm{d}x}{1-x^2}=\frac{1}{m}\frac{(l+m)!}{(l-m)!}\delta_{mm'};$$

证明如下. 由 P_l^m 和 $\mathrm{P}_{l'}^{m'}$ 所分别满足的微分方程得

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\left(\mathrm{P}_l^m\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{P}_{l'}^{m'}-\mathrm{P}_{l'}^{m'}\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{P}_l^m\right)\right]$$
$$=[l(l+1)-l'(l'+1)]\mathrm{P}_l^m\mathrm{P}_{l'}^{m'}-\frac{m^2-m'^2}{1-x^2}\mathrm{P}_l^m\mathrm{P}_{l'}^{m'}.$$

求其积分,左边的积分为零,结果是

$$[l(l+1)-l'(l'+1)]\int_{-1}^{1}\mathrm{P}_l^m\mathrm{P}_{l'}^{m'}\mathrm{d}x=(m^2-m'^2)\int_{-1}^{1}\frac{\mathrm{P}_l^m\mathrm{P}_{l'}^{m'}}{1-x^2}\mathrm{d}x;$$

因此,正交性得证,即

$$\int_{-1}^{1}\mathrm{P}_l^m\mathrm{P}_{l'}^m\mathrm{d}x=0,\qquad 若\ l\neq l',$$

$$\int_{-1}^{1}\frac{\mathrm{P}_l^m\mathrm{P}_l^{m'}}{1-x^2}\mathrm{d}x=0,\quad 若\ m\neq m'.$$

下面来证明

$$\int_{-1}^{1}[\mathrm{P}_l^m(x)]^2\mathrm{d}x=\frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}.$$

从递推关系

$$(2l+1)x\mathrm{P}_l^m=(l+m)\mathrm{P}_{l-1}^m+(l-m+1)\mathrm{P}_{l+1}^m$$

出发,分别乘以 P_{l+1}^m 和 P_{l-1}^m 并积分得到

$$(l-m+1)\int_{-1}^{1}[\mathrm{P}_{l+1}^m]^2\mathrm{d}x=(2l+1)\int_{-1}^{1}x\mathrm{P}_{l+1}^m\mathrm{P}_l^m\mathrm{d}x,$$

$$(l+m)\int_{-1}^{1}[\mathrm{P}_{l-1}^m]^2\mathrm{d}x=(2l+1)\int_{-1}^{1}x\mathrm{P}_l^m\mathrm{P}_{l-1}^m\mathrm{d}x;$$

这里已利用了正交关系. 在前式中令 $l\to l-1$,并利用后式,可求得

$$(2l+1)(l-m)\int_{-1}^{1}[\mathrm{P}_l^m]^2\mathrm{d}x=(2l-1)(l+m)\int_{-1}^{1}[\mathrm{P}_{l-1}^m]^2\mathrm{d}x;$$

由此递推,容易得到

$$\int_{-1}^{1}[\mathrm{P}_l^m]^2\mathrm{d}x=\frac{2m+1}{2l+1}\frac{(l+m)!}{(l-m)!(2m)!}\int_{-1}^{1}[\mathrm{P}_m^m]^2\mathrm{d}x.$$

---

① 王竹溪,郭敦仁:《特殊函数概论》,5.12 节.

② Bateman, *Partial differential Equations of Mathematical Physics*, pp. 362—364.

因为

$$(-)^m \mathrm{P}_m^m(x) = (1-x^2)^{\frac{m}{2}} \frac{\mathrm{d}^m}{\mathrm{d}x^m}\mathrm{P}_m(x) = \frac{(2m)!}{2^m(m!)^2} m!(1-x^2)^{\frac{m}{2}}$$

$$= (2m-1)!!(1-x^2)^{\frac{m}{2}},$$

这里$(2m-1)!! = 1 \cdot 3 \cdots (2m-1)$，于是①

$$\int_{-1}^{1}[\mathrm{P}_m^m(x)]^2 \mathrm{d}x = [(2m-1)!!]^2 \int_{-1}^{1}(1-x^2)^m \mathrm{d}x = \frac{2}{2m+1}(2m)!;$$

代入后即得欲证明之关系.

再来证明

$$\int_{-1}^{1}[\mathrm{P}_l^m(x)] \frac{\mathrm{d}x}{1-x^2}\mathrm{d}x = \frac{1}{m}\frac{(l+m)!}{(l-m)!}.$$

从递推关系

$$(1-x^2)\frac{\mathrm{d}\mathrm{P}_l^m}{\mathrm{d}x} = (l+m)\mathrm{P}_{l-1}^m - lx\mathrm{P}_l^m,$$

$$(1-x^2)\frac{\mathrm{d}\mathrm{P}_{l-1}^m}{\mathrm{d}x} = lx\mathrm{P}_{l-1}^m - (l-m)\mathrm{P}_l^m$$

出发，分别乘以$(1-x^2)^{-1}\mathrm{P}_{l-1}^m$和$(1-x^2)^{-1}\mathrm{P}_l^m$后相加，再积分，可以得到

$$(l+m)\int_{-1}^{1}[\mathrm{P}_{l-1}^m(x)]^2 \frac{\mathrm{d}x}{1-x^2} - (l-m)\int_{-1}^{1}[\mathrm{P}_l^m(x)]^2 \frac{\mathrm{d}x}{1-x^2}$$

$$= \int_{-1}^{1}\frac{\mathrm{d}}{\mathrm{d}x}[\mathrm{P}_l^m \mathrm{P}_{l-1}^m]\mathrm{d}x = 0 \quad (\text{若 } m > 0).$$

于是，利用前面的结果，可以得到

$$\int_{-1}^{1}[\mathrm{P}_l^m(x)]^2 \frac{\mathrm{d}x}{1-x^2}$$

$$= \frac{(l+m)!}{(l-m)!(2m)!}\int_{-1}^{1}[\mathrm{P}_m^m(x)]^2 \frac{\mathrm{d}x}{1-x^2}$$

---

① 这里应用了下列积分结果

$$\int_{-1}^{1}(1-x^2)^n \mathrm{d}x = 2\frac{(2n)!!}{(2n+1)!!},$$

推导如下. 首先，令 $x = \cos\theta$，

$$\int_{-1}^{1}(1-x^2)^n \mathrm{d}x = 2\int_{0}^{1}(1-x^2)^n \mathrm{d}x = 2\int_{0}^{\frac{\pi}{2}}\sin^{2n+1}\theta \mathrm{d}\theta;$$

再用分部积分

$$\int_{0}^{\frac{\pi}{2}}\sin^{2n+1}\theta \mathrm{d}\theta = (-\cos\theta\sin^{2n}\theta)\Big|_{0}^{\frac{\pi}{2}} + 2n\int_{0}^{\frac{\pi}{2}}\cos^2\theta\sin^{2n-1}\theta \mathrm{d}\theta,$$

于是，

$$2\int_{0}^{\frac{\pi}{2}}\sin^{2n+1}\theta \mathrm{d}\theta = 2\frac{2n}{2n+1}\int_{0}^{\frac{\pi}{2}}\sin^{2n-1}\theta \mathrm{d}\theta = 2\frac{(2n)!!}{(2n+1)!!}\int_{0}^{\frac{\pi}{2}}\sin\theta \mathrm{d}\theta = 2\frac{(2n)!!}{(2n+1)!!}.$$

$$= \frac{(l+m)!}{(l-m)!(2m)!}[(2m-1)!!]^2\int_{-1}^{1}(1-x^2)^{m-1}\mathrm{d}x$$
$$= \frac{1}{m}\cdot\frac{(l+m)!}{(l-m)!}.$$

得证.

顺便指出,令 $m=0$ 时得 $\mathrm{P}_l(x)$的正交归一关系

$$\int_{-1}^{1}\mathrm{P}_l(x)\mathrm{P}_{l'}(x)\mathrm{d}x = \frac{2}{2l+1}\delta_{ll'}.$$

同时应该指出,$\{\mathrm{P}_l(x)\}$和$\{\mathrm{P}_l^m(x)\}(l\geqslant m)$分别构成区间$[-1,1]$中的完备的正交函数系.任意一个在区间$[-1,1]$中连续且在端点为零的函数 $f(x)$可以用任意阶(m)的连带 Legendre 函数$\mathrm{P}_l^m(x)$在平均收敛的意义上展开为

$$f(x) = \sum_{l\geqslant m}^{\infty}a_l\mathrm{P}_l^m(x),$$

其中

$$a_l = \frac{2l+1}{2}\frac{(l-m)!}{(l+m)!}\int_{-1}^{1}f(x)\mathrm{P}_l^m(x)\mathrm{d}x.$$

12.5 非正则奇点邻域内的常规解

对于二阶线性常微分方程

$$\frac{\mathrm{d}^2u}{\mathrm{d}z^2}+p(z)\frac{\mathrm{d}u}{\mathrm{d}z}+q(z)u = 0$$

的求解,首先看欲求解区域的性质:若在常点邻域则可采用 Taylor 级数解,若在正则奇点邻域则可求其正则解.现在来考虑非此两类情况下的另一类解法.

12.5.1 方程的非正则奇点

若一点 $z=z_{\mathrm{ir}}$既非方程的常点,又非其正则奇点,则称为方程的**非正则奇点**.

例如,微分方程

$$\frac{\mathrm{d}^2u}{\mathrm{d}z^2}+\left(a+\frac{b}{z}+\frac{c}{z^2}\right)u = 0,$$

$$p(z) = 0,\quad q(z) = a+\frac{b}{z}+\frac{c}{z^2};$$

$z=0$ 是 $q(z)$的二阶极点,所以是方程的正则奇点.再看 $z=\infty$点,经 $z_1=\dfrac{1}{z}$变换有

$$p_1(z_1) = \frac{2}{z_1}-\frac{1}{z_1^2}p\left(\frac{1}{z_1}\right) = \frac{2}{z_1},$$

$$q_1(z_1)=\frac{1}{z_1^4}q\left(\frac{1}{z_1}\right)=\frac{a}{z_1^4}+\frac{b}{z_1^3}+\frac{c}{z_1^2},$$

$z_1=0$ 是 $p_1(z_1)$的单极点，是 $q_1(z_1)$的四阶($a\neq0$)或三阶($a=0,b\neq0$)极点；所以，$z=\infty$是方程的非正则奇点；只有在 $a=b=0$ 时才是正则奇点.

12.5.2 常规解

若 $z_{ir}=\infty$是方程的一个非正则奇点，则可试用下列形式的**常规解**：

$$u=e^{Q(z)}z^{\rho}\sum_{n=0}^{\infty}a_nz^{-n},$$

其中 $Q(z)$是 z 的多项式.

首先，以简单例子来阐明此常规解想法的由来.

考虑下列一阶微分方程

$$\frac{du}{dz}+\left(\frac{a}{z}+\frac{b}{z^2}+\frac{c}{z^3}\right)u=0,$$

$z_{ir}=0$ 是方程的非正则奇点. 此方程很容易解出如下：

$$\frac{du}{u}=-\left(\frac{a}{z}+\frac{b}{z^2}+\frac{c}{z^3}\right)dz,$$

$$\ln u=-a\ln z+\frac{b}{z}+\frac{c}{2z^2}+\text{const},$$

$$u=Cz^{-a}e^{\left(bz^{-1}+\frac{c}{2}z^{-2}\right)}.$$

这才引起上述常规解的想法.

现在考虑一般方程

$$\frac{d^2u}{dz^2}+p(z)\frac{du}{dz}+q(z)u=0,$$

假定它在 $z_{ir}=\infty$邻域内有常规解. 若令 $u=e^{Q(z)}v(z)$，代入上述方程后得出 $v(z)$的微分方程

$$\frac{d^2v}{dz^2}+p^*(z)\frac{dv}{dz}+q^*(z)v=0,$$

$$p^*(z)=p(z)+2Q'(z),$$

$$q^*(z)=q(z)+p(z)Q'(z)+Q''(z)+[Q'(z)]^2;$$

则要求 $v(z)$应具有正则形式解 $v(z)=z^{\rho}\sum_{n=0}^{\infty}a_nz^{-n}$. 将此形式解代入后一方程，经过分析可得下面的**定理**：在极点型非正则奇点邻域内存在一个正则形式解的必要条件是 $p^*(z)$的阶必须大于 $q^*(z)$的阶. 这里函数的阶指的是其展开式中 z 最高次幂的指数. 这个定理中的条件，可用来确定 $Q(z)$的具体形式，以及从而确定 ρ 和 a_n 等.

应该指出，对于方程在 $z_{ir}=\infty$ 邻域内的常规解

$$u = e^{Q(z)} z^{\rho} \sum_{n=0}^{\infty} a_n z^{-n},$$

z 的多项式 $Q(z)$ 是决定性因素，它描述此解 $u(z)$ 的本质奇点（$z_{ir}=\infty$）行为.

总之，对于方程的非正则邻域内的求解，首先试用常规解；如果没有常规解，再试别的办法.

12.5.3 汇合型超几何方程①②

下面将以汇合型超几何方程

$$z \frac{d^2 u}{dz^2} + (\gamma - z) \frac{du}{dz} - \alpha u = 0$$

作为具体例子，进一步阐明求常规解的步骤.

在此首先指出，若在超几何方程

$$z(1-z) \frac{d^2 u}{dz^2} + [\gamma - (\alpha + \beta + 1) z] \frac{du}{dz} - \alpha\beta u = 0$$

中作变换 $z \to \frac{z}{b}$ 后用 b 除，再令 $b=\beta \to \infty$ 即可得上述方程. 在此过程中，方程的奇点则由 $0,1,\infty \to 0,b,\infty \to 0,\infty$；其中 0 仍为正则奇点，而 ∞ 乃原来两正则奇点之汇合，变为非正则奇点；故将所得方程称为**汇合型超几何方程**.

(1) $z_r=0$ 邻域内的正则解. 汇合型超几何函数

汇合型超几何方程

$$z \frac{d^2 u}{dz^2} + (\gamma - z) \frac{du}{dz} - \alpha u = 0$$

在正则奇点 $z_r=0$ 邻域内的两个正则解容易求得为

$$u_1 = F(\alpha;\gamma;z),$$
$$u_2 = z^{1-\gamma} F(\alpha - \gamma + 1; 2 - \gamma; z),$$

其中

$$F(\alpha;\gamma;z) = \sum_{n=0}^{\infty} \frac{(\alpha)_n}{n!(\gamma)_n} z^n \quad (\gamma \neq 0 \text{ 或负整数}),$$

称为**汇合型超几何函数**；它也可通过上述同样变换过程从超几何函数得到.

(2) $z_{ir}=\infty$ 邻域内的常规解

现在来求在 $z_{ir}=\infty$ 邻域内的常规解

$$u = e^{Q(z)} v(z), \quad v(z) = z^{\rho} \sum_{n=0}^{\infty} a_n z^{-n}.$$

---

① Whittaker, Watson, *Modern Analysis*, Ch. XVI.

② 王竹溪，郭敦仁：《特殊函数概论》，第六章.

代入原方程得到

$$\frac{\mathrm{d}^2 v}{\mathrm{d}z^2} + p^*(z)\frac{\mathrm{d}v}{\mathrm{d}z} + q^*(z)v = 0,$$

$$p^*(z) = \frac{\gamma}{z} - 1 + 2Q'(z),$$

$$q^*(z) = -\frac{\alpha}{z} + \left(\frac{\gamma}{z} - 1\right)Q'(z) + Q''(z) + [Q'(z)]^2.$$

此方程有正则形式解的条件要求 $p^*(z)$的阶大于 $q^*(z)$的阶，立刻看出 $Q'(z)$的方次不能超过 0，而且要求

$$Q'(z)[Q'(z) - 1] = 0$$

$$\to \begin{cases} Q'(z) = 0 \to Q(z) = (\text{const} \xlongequal{\text{令}})0, \\ Q'(z) = 1 \to Q(z) = z. \end{cases}$$

因此

$$u = v \to \overset{-1}{z\frac{\mathrm{d}^2 v}{\mathrm{d}z^2}} + (\overset{-1}{\gamma} - \overset{0}{z})\frac{\mathrm{d}v}{\mathrm{d}z} - \overset{0}{\alpha v} = 0,$$

$$u = \mathrm{e}^z v \to z\frac{\mathrm{d}^2 v}{\mathrm{d}z^2} + (\gamma + z)\frac{\mathrm{d}v}{\mathrm{d}z} + (\gamma - \alpha)v = 0.$$

下面方程中，若令 $z \to -z, \gamma - \alpha \to \alpha'$，即与上面方程同为汇合型超几何方程. 所以，只需解上面的方程即可. 用 $v = z^\rho \sum_{n=0}^{\infty} a_n z^{-n}$ 代入，注意到这是二阶步方程，容易得到（注意按**降幂序**）

$$-a_0(\rho + \alpha)z^\rho - \sum_{n=1}^{\infty} a_n\{(\rho - n) + \alpha\}z^{\rho - n}$$
$$+ \sum_{n=1}^{\infty} a_{n-1}\{(\rho - n + 1)(\rho - n) + \gamma(\rho - n + 1)\}z^{\rho - n} = 0.$$

z^ρ 的系数$(\rho + \alpha) = 0$ 给出 $\rho = -\alpha$（令 $a_0 = 1$），$z^{\rho - n}$的系数为零给出递推关系（用 $\rho = -\alpha$ 代入）

$$a_n = -\frac{(\alpha - 1 + n)(\alpha - \gamma + n)}{n}a_{n-1}.$$

由此得（令 $a_0 = 1$）

$$a_n = (-)^n \frac{(\alpha)_n(\alpha - \gamma + 1)_n}{n!} \quad (n = 0, 1, 2, \cdots);$$

于是（下面“$\simeq$”表示“渐近等于”，其意义在 12.5.4 节之(2)后有说明），

$$u_3 \simeq z^{-\alpha}\sum_{n=0}^{\infty}(-)^n \frac{(\alpha)_n(\alpha - \gamma + 1)_n}{n!}z^{-n}.$$

同时令 $z \to -z$，$\alpha \to \gamma - \alpha$，容易得到

$$u_4 \simeq \mathrm{e}^z(-z)^{\alpha-\gamma}\sum_{n=0}^{\infty}(-)^n\frac{(\gamma-\alpha)_n(1-\alpha)_n}{n!}(-z)^{-n}.$$

为了得到更加对称的形式，可令

$$\alpha=\frac{1}{2}+m-k,\quad \gamma=1+2m;$$

因而得到

$$\begin{aligned}
u_3 &\simeq z^{-\left[m-\left(k-\frac{1}{2}\right)\right]}\sum_{n=0}^{\infty}(-)^n\frac{\left[m-\left(k-\frac{1}{2}\right)\right]_n\left[-m-\left(k-\frac{1}{2}\right)\right]_n}{n!}z^{-n}\\
&\simeq z^{-\left[m-\left(k-\frac{1}{2}\right)\right]}\left\{1+\sum_{n=1}^{\infty}\frac{1}{n!}\left[m^2-\left(k-\frac{1}{2}\right)^2\right]\cdots\right.\\
&\qquad\left.\cdot\left[m^2-\left(k-n+\frac{1}{2}\right)^2\right]z^{-n}\right\},\\
u_4 &\simeq \mathrm{e}^z(-z)^{-\left[m+\left(k+\frac{1}{2}\right)\right]}\sum_{n=0}^{\infty}(-)^n\frac{1}{n!}\left[m+\left(k+\frac{1}{2}\right)\right]_n\\
&\qquad\cdot\left[-m+\left(k+\frac{1}{2}\right)\right]_n(-z)^{-n}\\
&\simeq \mathrm{e}^z(-z)^{-\left[m+\left(k+\frac{1}{2}\right)\right]}\left\{1+\sum_{n=1}^{\infty}\frac{1}{n!}\left[m^2-\left(k+\frac{1}{2}\right)^2\right]\cdots\right.\\
&\qquad\left.\cdot\left[m^2-\left(k+n-\frac{1}{2}\right)^2\right](-z)^{-n}\right\}.
\end{aligned}$$

12.5.4　Whittaker 方程

若在汇合型超几何方程

$$z\frac{\mathrm{d}^2u}{\mathrm{d}z^2}+(\gamma-z)\frac{\mathrm{d}u}{\mathrm{d}z}-\alpha u=0$$

中令 $u=\mathrm{e}^{z/2}z^{-\gamma/2}w(z)$，得到

$$\frac{\mathrm{d}^2w}{\mathrm{d}z^2}+\left[-\frac{1}{4}+\left(\frac{\gamma}{2}-\alpha\right)\frac{1}{z}+\frac{\gamma}{2}\left(1-\frac{\gamma}{2}\right)\frac{1}{z^2}\right]w=0,$$

这就是在 12.5.1 节中曾经引用过的形式(所用参数不同). 若再作上述参数变换 $\alpha=\frac{1}{2}+m-k,\gamma=1+2m$，则得 **Whittaker 方程**：

$$\overset{-2}{\frac{\mathrm{d}^2w}{\mathrm{d}z^2}}+\left[\overset{0}{-\frac{1}{4}}+\overset{-1}{\frac{k}{z}}+\overset{-2}{\frac{\frac{1}{4}-m^2}{z^2}}\right]w=0.$$

上述微分方程含三个阶步，求递推公式很不方便，要用特殊方法；如令 $\overline{w}=\mathrm{e}^{-\frac{1}{2}z}w$，

就可得到两个阶步的方程

$$\frac{\mathrm{d}^2\overline{w}}{\mathrm{d}z^2}+\frac{\mathrm{d}\overline{w}}{\mathrm{d}z}+\left[\frac{k}{z}+\frac{\frac{1}{4}-m^2}{z^2}\right]\overline{w}=0.$$

也就可以按通常办法求解.

然而由于前面已求出汇合型超几何方程的解,因此按 $w(z)=\mathrm{e}^{-z/2}z^{\gamma/2}u(z)$ 并作参数变换,就可直接得到 Whittaker 方程的解.

(1) Whittaker M 函数

Whittaker 方程在正则奇点 $z_{\mathrm{r}}=0$ 邻域内的正则解是 **Whittaker M 函数**:(若 $2m$ 不是整数时,它们线性无关)

$$\mathrm{M}_{k,m}(z)=\mathrm{e}^{-z/2}z^{\left(\frac{1}{2}+m\right)}\mathrm{F}\left(\frac{1}{2}+m-k;1+2m;z\right),$$

$$\mathrm{M}_{k,-m}(z)=\mathrm{e}^{-z/2}z^{\left(\frac{1}{2}-m\right)}\mathrm{F}\left(\frac{1}{2}-m-k;1-2m;z\right);$$

其中 $\mathrm{F}(\alpha;\gamma;z)$是前面引进过的汇合型超几何函数

$$\mathrm{F}(\alpha;\gamma;z)=\sum_{n=0}^{\infty}\frac{(\alpha)_n}{n!(\gamma)_n}z^n;$$

$$(\lambda)_0=1,\quad(\lambda)_n=\lambda(\lambda+1)\cdots(\lambda+n-1).$$

(2) Whittaker W 函数

Whittaker 方程在非正则奇点 $z_{\mathrm{ir}}=\infty$邻域内的常规解是 **Whittaker W 函数**:

$$\mathrm{W}_{k,m}(z)\simeq\mathrm{e}^{-z/2}z^k\left\{1+\sum_{n=1}^{\infty}\frac{1}{n!}\left[m^2-\left(k-\frac{1}{2}\right)^2\right]\right.$$
$$\left.\cdots\left[m^2-\left(k-n+\frac{1}{2}\right)^2\right]z^{-n}\right\},$$

$$\mathrm{W}_{-k,m}(z)\simeq\mathrm{e}^{z/2}(-z)^{-k}\left\{1+\sum_{n=1}^{\infty}\frac{1}{n!}\left[m^2-\left(k+\frac{1}{2}\right)^2\right]\right.$$
$$\left.\cdots\left[m^2-\left(k+n-\frac{1}{2}\right)^2\right](-z)^{-n}\right\}.$$

应该指出,在非正则奇点处,正则形式解中的级数并不总是收敛的,它实际上只是解在 $z\to\infty$时的**渐近展开**. 所以,在 $\mathrm{W}_{\pm k,m}$以及 u_3,u_4 的表达式中用渐近等于的符号"$\simeq$"表示. 用渐近展开,以取适当项数为宜,因为近似程度一般不能通过多取项数而无限制地改善.

与求正则解时的情况相类似,对于给定边界条件,如要求解在 0,∞处为有限,同样会导致所谓本征值问题.

(3) 解析延拓问题

另外,关于 Whittaker 函数的延拓问题,给出其结果如下.

$$W_{k,m}(z)=\frac{\Gamma(-2m)}{\Gamma\left(\frac{1}{2}-m-k\right)}M_{k,m}(z)+\frac{\Gamma(2m)}{\Gamma\left(\frac{1}{2}+m-k\right)}M_{k,-m}(z)$$

$$(2m\neq 整数,\quad |\arg z|<3\pi/2),$$

$$W_{-k,m}(-z)=\frac{\Gamma(-2m)}{\Gamma\left(\frac{1}{2}-m+k\right)}M_{-k,m}(-z)+\frac{\Gamma(2m)}{\Gamma\left(\frac{1}{2}+m+k\right)}M_{-k,-m}(-z)$$

$$(2m\neq 整数,\quad |\arg(-z)|<3\pi/2);$$

$$M_{k,m}(z)=\frac{\Gamma(2m+1)}{\Gamma\left(\frac{1}{2}+m+k\right)}e^{\left(\frac{1}{2}+m-k\right)\pi i}W_{k,m}(z)+\frac{\Gamma(2m+1)}{\Gamma\left(\frac{1}{2}+m-k\right)}e^{-k\pi i}W_{-k,m}(ze^{-\pi i})$$

$$(2m\neq 负整数,\quad -\pi/2<\arg z<3\pi/2),$$

$$M_{k,m}(z)=\frac{\Gamma(2m+1)}{\Gamma\left(\frac{1}{2}+m+k\right)}e^{-\left(\frac{1}{2}+m-k\right)\pi i}W_{k,m}(z)+\frac{\Gamma(2m+1)}{\Gamma\left(\frac{1}{2}+m-k\right)}e^{k\pi i}W_{-k,m}(ze^{\pi i})$$

$$(2m\neq 负整数,\quad -3\pi/2<\arg z<\pi/2).$$

12.5.5 Bessel 方程①

(1) Bessel 方程

若在汇合型超几何方程

$$z\frac{d^2u}{dz^2}+(\gamma-z)\frac{du}{dz}-\alpha u=0$$

中令 $z=2i\zeta, u(2i\zeta)=\zeta^{-\nu}e^{i\zeta}\psi(\zeta)$,容易得到

$$\frac{d^2\psi}{d\zeta^2}+\frac{\gamma-2\nu}{\zeta}\frac{d\psi}{d\zeta}+\left[1+\frac{i(\gamma-2\alpha)}{\zeta}+\frac{\nu(\nu+1-\gamma)}{\zeta^2}\right]\psi=0;$$

特别是,若 $\gamma=2\alpha$ 以及 $\gamma=2\nu+1$ 或 $\gamma=2(\nu+1)$,则化为 **Bessel 方程**(ν 称为方程的阶)

$$\frac{d^2\psi}{d\zeta^2}+\frac{1}{\zeta}\frac{d\psi}{d\zeta}+\left(1-\frac{\nu^2}{\zeta^2}\right)\psi=0,$$

或**球 Bessel 方程**(令 $\nu\to l$)

$$\frac{d^2\psi}{d\zeta^2}+\frac{2}{\zeta}\frac{d\psi}{d\zeta}+\left[1-\frac{l(l+1)}{\zeta^2}\right]\psi=0.$$

可以看出,Bessel 方程和球 Bessel 方程分别是在柱面坐标下和球面坐标下用

---

① 王竹溪、郭敦仁:《特殊函数概论》;7.1—7.9 节.

分离变量法求解波动方程时所得的径向方程[参见 11.3.4 之(3)和(2)].

还应指出，通过令 $\psi(\zeta)=\zeta^{-\frac{1}{2}}\bar{\psi}(\zeta)$，球 Bessel 方程可化为半奇数阶 $\left(l+\frac{1}{2}\right)$ 的 Bessel 方程. 另一方面，若令 $\psi(\zeta)=\text{const}\cdot\zeta^{-\frac{1}{2}}\cdot w(2\mathrm{i}\zeta)$，$z=2\mathrm{i}\zeta$，则 Bessel 方程化为 Whittaker 方程($k=0,m=\nu$)

$$\frac{\mathrm{d}^2 w}{\mathrm{d}z^2}+\left[-\frac{1}{4}+\frac{\frac{1}{4}-\nu^2}{z^2}\right]w=0.$$

最后，若在 Bessel 方程中令 $\zeta=\mathrm{i}x$(x 是实数)，所得方程

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}+\frac{1}{x}\frac{\mathrm{d}\psi}{\mathrm{d}x}-\left(1+\frac{\nu^2}{x^2}\right)\psi=0,$$

则称为**变形(或虚宗量)Bessel 方程**.

(2) Bessel 函数

按照上一小节的讨论，得到 Bessel 方程的一个解是 Bessel 函数

$$\begin{aligned}\mathrm{J}_\nu(\zeta)&=\frac{1}{2^\nu\Gamma(\nu+1)}\zeta^\nu\mathrm{e}^{-\mathrm{i}\zeta}\mathrm{F}\left(\nu+\frac{1}{2};2\nu+1;2\mathrm{i}\zeta\right)\\&=\frac{(2\mathrm{i})^{-\left(\nu+\frac{1}{2}\right)}}{2^\nu\Gamma(\nu+1)}\zeta^{-\frac{1}{2}}\mathrm{M}_{0,\nu}(2\mathrm{i}\zeta)\\&=\sum_{k=0}^{\infty}(-)^k\frac{1}{k!\Gamma(\nu+k+1)}\left(\frac{\zeta}{2}\right)^{2k+\nu}.\end{aligned}$$

最后一行的展开式可从直接求 Bessel 方程在 $z_r=0$ 邻域内的正则解得到，因为系数的递推关系是

$$a_{2k}=-\frac{a_{2k-2}}{2k(2\nu+2k)}=(-)^k\frac{\Gamma(\nu+1)}{2^{2k}k!\Gamma(\nu+k+1)}a_0,$$

这里取 $a_0=[2^\nu\Gamma(\nu+1)]^{-1}$.

应该注意，当 2ν 不是整数时，$\mathrm{J}_{-\nu}(\zeta)$ 是与 $\mathrm{J}_\nu(\zeta)$ 线性无关的另一个解. $\mathrm{J}_{\pm\nu}(\zeta)$ 称为 **ν 阶第一类 Bessel 函数**，简称 **ν 阶 Bessel 函数**.

当 $\nu=n(n=0,1,2,\cdots)$ 时，因 $\Gamma(-n+k+1)\to\infty(k<n)$，容易得到

$$\mathrm{J}_{-n}(\zeta)=(-)^n\mathrm{J}_n(\zeta),$$

$\mathrm{J}_{-n}(\zeta)$ 与 $\mathrm{J}_n(\zeta)$ 是线性相关的.

通常采用

$$\mathrm{N}_\nu(\zeta)=\frac{\cos\nu\pi\mathrm{J}_\nu(\zeta)-\mathrm{J}_{-\nu}(\zeta)}{\sin\nu\pi}$$

作为 ν 阶 Bessel 方程的第二解，称为**第二类 Bessel 函数**，亦称 **Neumann 函数**. 当 $\nu=n$(整数)时，右边为不定式，其极限存在，而且是 Bessel 方程的解，与 $\mathrm{J}_n(\zeta)$ 线性

无关.

还可引进**第三类 Bessel 函数**

$$\mathrm{H}_\nu^{(1)}(\zeta)=\mathrm{J}_\nu(\zeta)+\mathrm{i}\mathrm{N}_\nu(\zeta),$$
$$\mathrm{H}_\nu^{(2)}(\zeta)=\mathrm{J}_\nu(\zeta)-\mathrm{i}\mathrm{N}_\nu(\zeta);$$

亦称 **Hankel 函数**.

显而易见,对于任何 ν(包括 ν 等于整数 n),$\mathrm{J}_\nu(\zeta)$,$\mathrm{N}_\nu(\zeta)$,$\mathrm{H}_\nu^{(1)}(\zeta)$,$\mathrm{H}_\nu^{(2)}(\zeta)$中的任意两个总是线性无关的.

容易证明,上述 Bessel 函数中的任何一个,用 $\mathrm{Z}_\nu(\zeta)$表示,都满足下列递推关系:

$$\frac{\mathrm{d}}{\mathrm{d}\zeta}(\zeta^\nu \mathrm{Z}_\nu)=\zeta^\nu \mathrm{Z}_{\nu-1},$$
$$\frac{\mathrm{d}}{\mathrm{d}\zeta}(\zeta^{-\nu} \mathrm{Z}_\nu)=-\zeta^{-\nu} \mathrm{Z}_{\nu+1};$$

或者等价地,

$$\mathrm{Z}_{\nu-1}+\mathrm{Z}_{\nu+1}=\frac{2\nu}{\zeta}\mathrm{Z}_\nu,$$
$$\mathrm{Z}_{\nu-1}-\mathrm{Z}_{\nu+1}=2\frac{\mathrm{d}}{\mathrm{d}\zeta}\mathrm{Z}_\nu.$$

凡满足上述递推关系的函数统称为**柱函数**.

(3) 球 Bessel 函数

首先讨论一下**半奇数阶 Bessel 函数** $\mathrm{J}_{l+\frac{1}{2}}(l=0,\pm1,\pm2,\cdots)$的情形,它的一个重要特点是可用初等函数表达.例如,

$$\mathrm{J}_{\frac{1}{2}}(\zeta)=\sum_{k=0}^{\infty}\frac{(-)^k}{k!}\frac{1}{\Gamma\left(k+\frac{3}{2}\right)}\left(\frac{\zeta}{2}\right)^{2k+\frac{1}{2}}.$$

注意到 $\Gamma(z+1)=z\Gamma(z)$和 $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$,故

$$\begin{aligned}\Gamma\left(k+\frac{3}{2}\right)&=\left(k+\frac{1}{2}\right)\left(k-\frac{1}{2}\right)\cdots\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)\\&=\frac{1}{2^{2k+1}}\cdot\frac{(2k+1)!}{k!}\sqrt{\pi},\end{aligned}$$

$$\mathrm{J}_{\frac{1}{2}}(\zeta)=\sqrt{\frac{2}{\pi\zeta}}\sum_{k=0}^{\infty}\frac{(-)^k}{(2k+1)!}\zeta^{2k+1}=\sqrt{\frac{2}{\pi\zeta}}\sin\zeta.$$

类似地,可以证明

$$\mathrm{J}_{-\frac{1}{2}}(\zeta)=\sqrt{\frac{2}{\pi\zeta}}\cos\zeta.$$

另一方面,由上一段(2)中前两个递推关系容易得到

$$\left(\frac{\mathrm{d}}{\zeta\mathrm{d}\zeta}\right)^m(\zeta^\nu Z_\nu)=\zeta^{\nu-m}Z_{\nu-m},$$

$$\left(\frac{\mathrm{d}}{\zeta\mathrm{d}\zeta}\right)^m(\zeta^{-\nu}Z_\nu)=(-)^m\zeta^{-\nu-m}Z_{\nu+m}.$$

于是,可以求出普遍的半奇数阶 Bessel 函数为

$$\begin{aligned}J_{l+\frac{1}{2}}(\zeta)&=(-)^l\zeta^{l+\frac{1}{2}}\left(\frac{\mathrm{d}}{\zeta\mathrm{d}\zeta}\right)^l\{\zeta^{-\frac{1}{2}}J_{\frac{1}{2}}(\zeta)\}\\&=(-)^l\sqrt{\frac{2}{\pi\zeta}}\zeta^{l+1}\left(\frac{\mathrm{d}}{\zeta\mathrm{d}\zeta}\right)^l\left\{\frac{\sin\zeta}{\zeta}\right\}\quad(l=0,1,2,\cdots),\end{aligned}$$

$$J_{-l-\frac{1}{2}}(\zeta)=\sqrt{\frac{2}{\pi\zeta}}\zeta^{l+1}\left(\frac{\mathrm{d}}{\zeta\mathrm{d}\zeta}\right)^l\left\{\frac{\cos\zeta}{\zeta}\right\}\quad(l=0,1,2,\cdots).$$

于是,根据前面(1)段中曾经指出过的,球 Bessel 方程与半奇数阶 Bessel 方程的关系,可以定义**球 Bessel 函数**如下

$$\psi_l(\zeta)=\sqrt{\frac{\pi}{2\zeta}}\Psi_{l+\frac{1}{2}}(\zeta),$$

其中 $\Psi_{l+\frac{1}{2}}$ 表示 $J_{l+\frac{1}{2}}$, $N_{l+\frac{1}{2}}$, $H^{(1)}_{l+\frac{1}{2}}$, $H^{(2)}_{l+\frac{1}{2}}$ 中任何一个,而 ψ_l 则表示相应的球 Bessel 函数 j_l, n_l, $h_l^{(1)}$, $h_l^{(2)}$;现在 l 可以是任意的.

球 Bessel 函数满足下列基本递推关系:

$$\psi_{l-1}+\psi_{l+1}=\frac{2l+1}{\zeta}\psi_l,$$

$$l\psi_{l-1}-(l+1)\psi_{l+1}=(2l+1)\frac{\mathrm{d}\psi_l}{\mathrm{d}\zeta}.$$

(4) Bessel 函数的积分表示

从 Bessel 函数 $J_\nu(\zeta)$ 的级数表示

$$J_\nu(\zeta)=\left(\frac{\zeta}{2}\right)^\nu\sum_{k=0}^{\infty}\frac{(-)^k}{k!}\frac{1}{\Gamma(\nu+k+1)}\left(\frac{\zeta}{2}\right)^{2k}$$

出发,用 $\Gamma(\nu+k+1)$ 的周线积分公式(见 12.2.6 节之(3))

$$\frac{1}{\Gamma(\nu+k+1)}=\frac{1}{2\pi\mathrm{i}}\int_{-\infty}^{(0+)}\mathrm{e}^t t^{-\nu-k-1}\mathrm{d}t\quad(|\arg t|<\pi)$$

代入,交换求积分与求和次序

$$\begin{aligned}J_\nu(\zeta)&=\left(\frac{\zeta}{2}\right)^\nu\sum_{k=0}^{\infty}\frac{(-)^k}{k!}\left(\frac{\zeta}{2}\right)^{2k}\frac{1}{2\pi\mathrm{i}}\int_{-\infty}^{(0+)}t^{-\nu-k-1}\mathrm{e}^t\mathrm{d}t\\&=\left(\frac{\zeta}{2}\right)^\nu\frac{1}{2\pi\mathrm{i}}\int_{-\infty}^{(0+)}\left[\sum_{k=0}^{\infty}\frac{1}{k!}\left(-\frac{\zeta^2}{4t}\right)^k\right]t^{-\nu-1}\mathrm{e}^t\mathrm{d}t\quad(|\arg t|<\pi),\end{aligned}$$

可得下列积分表达式:

$$J_\nu(\zeta)=\frac{(\zeta/2)^\nu}{2\pi\mathrm{i}}\int_{-\infty}^{(0+)}\mathrm{e}^{t-\frac{\zeta^2}{4t}}t^{-\nu-1}\mathrm{d}t\quad(|\arg t|<\pi).$$

注意到，若$|\alpha|<\dfrac{\pi}{2}$，可将整个周线绕 $t=0$ 转一角度 α 而不改变积分值. 因此，将上式中的 t 换成 $\zeta t/2$，只要$|\arg\zeta|<\pi/2$，可得下列更简单的**积分表达式**：

$$\mathrm{J}_\nu(\zeta)=\frac{1}{2\pi\mathrm{i}}\int_{-\infty}^{(0+)}\mathrm{e}^{\frac{\zeta}{2}(t-t^{-1})}t^{-\nu-1}\mathrm{d}t\quad\left(|\arg\zeta|<\frac{\pi}{2},\ |\arg t|<\pi\right).$$

$-\infty$

若取如上图所示的积分周线，于是有

$$\mathrm{J}_\nu(\zeta)=\frac{1}{2\pi\mathrm{i}}\left\{(\mathrm{e}^{-\nu\pi\mathrm{i}}-\mathrm{e}^{\nu\pi\mathrm{i}})\int_1^{\infty}\mathrm{e}^{-\frac{\zeta}{2}(t-t^{-1})}t^{-\nu-1}\mathrm{d}t+\int^{(0+)}\mathrm{e}^{\frac{\zeta}{2}(t-t^{-1})}t^{-\nu-1}\mathrm{d}t\right\}.$$

前一积分中令 $t=\mathrm{e}^u$，后一积分中令 $t=\mathrm{e}^{\mathrm{i}\theta}$，变换次序，可得下列**积分表达式**：

$$\mathrm{J}_\nu(\zeta)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}(\zeta\sin\theta-\nu\theta)}\mathrm{d}\theta-\frac{\sin\nu\pi}{\pi}\int_0^{\infty}\mathrm{e}^{-\zeta\operatorname{sh}u-\nu u}\mathrm{d}u\quad(|\arg\zeta|<\pi/2),$$

或者（因为 $\sin(\zeta\sin\theta-\nu\theta)$ 是 θ 的奇函数其积分为零）

$$\mathrm{J}_\nu(\zeta)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\cos(\zeta\sin\theta-\nu\theta)\mathrm{d}\theta-\frac{\sin\nu\pi}{\pi}\int_0^{\infty}\mathrm{e}^{-\zeta\operatorname{sh}u-\nu u}\mathrm{d}u$$

$$(|\arg\zeta|<\pi/2).$$

特别是对于整数阶 Bessel 函数，有

$$\mathrm{J}_n(\zeta)=\frac{1}{2\pi\mathrm{i}}\int^{(0+)}\mathrm{e}^{\frac{\zeta}{2}(t-t^{-1})}t^{-n-1}\mathrm{d}t=\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}(\zeta\sin\theta-n\theta)}\mathrm{d}\theta$$

$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\cos(\zeta\sin\theta-n\theta)\mathrm{d}\theta\quad(n=0,\pm1,\pm2,\cdots).$$

(5) 展开公式

由上述 $\mathrm{J}_n(\zeta)$的第一个积分表达式立即可以看出，$\exp\left\{\dfrac{\zeta}{2}(t-t^{-1})\right\}$是 $\mathrm{J}_n(\zeta)$的**生成函数**，即其 **Laurent 展开**为

$$\mathrm{e}^{\frac{\zeta}{2}(t-t^{-1})}=\sum_{n=-\infty}^{\infty}\mathrm{J}_n(\zeta)t^n\quad(0<|t|<\infty).$$

在上式中若令 $t=\mathrm{i}\mathrm{e}^{\mathrm{i}\theta}$，$\zeta=kr$，并利用 $\mathrm{J}_{-n}=(-)^n\mathrm{J}_n$，可得

$$\begin{aligned}\mathrm{e}^{\mathrm{i}kr\cos\theta}&=\sum_{n=-\infty}^{\infty}\mathrm{J}_n(kr)\mathrm{i}^n\mathrm{e}^{\mathrm{i}n\theta}\\&=\mathrm{J}_0(kr)+\sum_{n=1}^{\infty}[\mathrm{J}_n(kr)\mathrm{i}^n\mathrm{e}^{\mathrm{i}n\theta}+\mathrm{J}_{-n}(kr)\mathrm{i}^{-n}\mathrm{e}^{-\mathrm{i}n\theta}]\\&=\mathrm{J}_0(kr)+2\sum_{n=1}^{\infty}\mathrm{i}^n\mathrm{J}_n(kr)\cos n\theta,\end{aligned}$$

它代表**平面波用柱面波展开的公式**.

下面再来推导平面波用 Legendre 多项式展开的公式. 设[见 12.4.4 节之(6)最后($m=0$)]

$$e^{ikr\cos\theta}=\sum_{l=0}^{\infty}a_l(kr)\mathrm{P}_l(\cos\theta),$$

$$a_l(kr)=\frac{2l+1}{2}\int_{-1}^{1}e^{ikrx}\mathrm{P}_l(x)\mathrm{d}x\quad(x=\cos\theta).$$

用 $\mathrm{P}_l(x)$的 Rodrigues 公式代入并分部积分 l 次可得

$$\begin{aligned}a_l(kr)&=\frac{2l+1}{2}\cdot\frac{1}{2^l l!}\int_{-1}^{1}e^{ikrx}\frac{\mathrm{d}^l}{\mathrm{d}x^l}(x^2-1)^l\mathrm{d}x\\&=\frac{2l+1}{2^{l+1}l!}\cdot(-)^l(ikr)^l\int_{-1}^{1}e^{ikrx}(x^2-1)^l\mathrm{d}x\\&=\frac{2l+1}{2^{l+1}l!}\cdot(-)^l\cdot(ikr)^l\int_{-1}^{1}\sum_{n=0}^{\infty}\frac{(ikr)^n}{n!}x^n(x^2-1)^l\mathrm{d}x\\&=\frac{2l+1}{2^{l+1}l!}\cdot(-)^l\cdot 2\sum_{s=0}^{\infty}\frac{(ikr)^{2s+l}}{(2s)!}\int_0^1 x^{2s}(x^2-1)^l\mathrm{d}x,\end{aligned}$$

最后一行的被积函数中已消除其积分值为零的奇函数部分. 对于最后一行的积分,令 $x^2=u$ 后可化为 B 函数,得到

$$\begin{aligned}\int_0^1 x^{2s}(x^2-1)^l\mathrm{d}x&=(-)^l\frac{1}{2}\int_0^1 u^{s-\frac{1}{2}}(1-u)^l\mathrm{d}u=(-)^l\frac{1}{2}\frac{\Gamma\left(s+\frac{1}{2}\right)\Gamma(l+1)}{\Gamma\left(l+s+\frac{3}{2}\right)}\\&=(-)^l\frac{l!(2s)!\sqrt{\pi}}{2^{2s+1}(s!)\Gamma\left(l+\frac{1}{2}+s+1\right)},\end{aligned}$$

这里应用了 $\Gamma\left(s+\frac{1}{2}\right)=\left(s-\frac{1}{2}\right)\left(s-\frac{3}{2}\right)\cdots\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)=\frac{(2s)!}{2^{2s}s!}\sqrt{\pi},\Gamma(l+1)=l!$.

于是,合并化简后得到

$$\begin{aligned}a_l(kr)&=(2l+1)i^l\sqrt{\frac{\pi}{2kr}}\sum_{s=0}^{\infty}(-)^s\frac{1}{s!}\frac{1}{\Gamma\left(l+\frac{1}{2}+s+1\right)}\left(\frac{kr}{2}\right)^{2s+l+\frac{1}{2}}\\&=(2l+1)i^l\sqrt{\frac{\pi}{2kr}}\mathrm{J}_{l+\frac{1}{2}}(kr)=(2l+1)i^l\mathrm{j}_l(kr),\end{aligned}$$

这里应用了 Bessel 函数的级数表示以及球 Bessel 函数的定义. 因此,最后得到**平面波用 Legendre 多项式展开的公式**:

$$e^{ikr\cos\theta}=\sum_{l=0}^{\infty}(2l+1)i^l\mathrm{j}_l(kr)\mathrm{P}_l(\cos\theta).$$

(6) 变形 Bessel 函数

由于变形 Bessel 方程可从 Bessel 方程令 $\zeta = \mathrm{i}x$ 而得，故可直接引进其两个线性无关解如下：

$$\begin{aligned}
\mathrm{I}_\nu(x) &= \mathrm{e}^{-\nu\pi\mathrm{i}/2}\mathrm{J}_\nu(x\mathrm{e}^{\pi\mathrm{i}/2}) \quad (-\pi < \arg x < \pi/2) \\
&= \mathrm{e}^{3\nu\pi\mathrm{i}/2}\mathrm{J}_\nu(x\mathrm{e}^{-3\pi\mathrm{i}/2}) \quad (\pi/2 < \arg x < \pi) \\
&= \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty} \frac{1}{k!}\frac{1}{\Gamma(\nu+k+1)}\left(\frac{x}{2}\right)^{2k},
\end{aligned}$$

$$\mathrm{K}_\nu(x) = \frac{\pi}{2\sin\nu\pi}[\mathrm{I}_{-\nu}(x) - \mathrm{I}_\nu(x)];$$

$\mathrm{I}_\nu(x)$和 $\mathrm{K}_\nu(x)$分别称为**变形第一类和第二类 Bessel 函数**. 注意到，当 x 是实数时，它们是实函数. 另外，当 $\nu = n$(整数)时，由 $\mathrm{J}_n = (-)^n\mathrm{J}_{-n}$及上面的定义，显然可见

$$\mathrm{I}_{-n}(x) = \mathrm{I}_n(x).$$

第 13 章　微分方程的数值解法

13.1　数值方法的重要性

现代科学技术问题最终常归结为求解微分方程. 在实际问题中我们所要求的并不是用一级数或积分所表达的函数，只要求出几点之值已经可以，并不需要那么准确的函数. 而且要求准确解在多数情况下非常麻烦、甚至根本不可能，不如用**数值方法**之既经济又实惠.

在实验里所求得的物理量之间的关系在方格纸上用几个点表示，要从这几个点找出某点的微分或某二点之间的积分，则不得不用数值方法.

数值方法还有一个好处是可以估计误差，因此可在我们所容许的误差范围内来选用适当的方法.

数值方法，由于其重要性，以及电子计算机的发展，将是数学中今后最有发展前途的一个分支.

13.2　Weierstrass 定理

对区间$[a,b]$上的任一连续实值函数 $f(x)$，存在多项式序列$\{P_n(x)\}$，它在区间$[a,b]$上一致收敛于 $f(x)$.

这是关于函数逼近的 **Weierstrass 定理**.

一个函数实际上是很多点子. Weierstrass 定理说，任何函数都可用有限项的幂级数来逼近，在小范围内可用很少几项的幂级数来逼近. 有时为了采用很少的项数来逼近，常将许多点子分成若干段，那么每一段就可用低次多项式来逼近了；这种方法称为分段逼近.

插值法是函数逼近的重要方法，也是导出数值微分、数值积分、微分方程数值解等其他许多数值方法的依据. 所以，下面先介绍一下插值方法.

13.3　插　值　法

插值法就是给定一组 $y=f(x)$值：(x_i,y_i)，$i=0,1,\cdots,n$，如何由此求出未给

各点的 $f(x)$ 值的数值方法.具体作法是先选定插值函数类型,再由插值条件来确定其中的待定参数.

常用的插值函数类型有:代数多项式型,三角多项式型,有理函数型.

三角插值一般广泛用于时间序列和周期现象的数值 Fourier 分析.在这方面,所谓"**快速 Fourier 变换**(FFT)"是尤其重要和成功的.

有理插值特别适用于被插函数具有极点或其他奇点的情况,例如,相变点邻近的热力学函数.有理插值还与 **Padé 逼近**密切相关,后者在许多物理问题中有着广泛的应用.

最有用的还是**多项式插值**,它既简单方便,又是多种常用数值微分和积分公式的依据.

下面仅对多项式插值作简要介绍.

13.3.1 多项式插值①②

对于给定数据组 (x_i, y_i), $i=0,1,\cdots,n$,取**插值多项式**为

$$P_n(x) = c_0 + c_1 x + \cdots + c_n x^n,$$

由**插值条件**

$$P_n(x_i) = y_i \quad (i = 0,1,\cdots,n)$$

来确定参数 $c_0, c_1, \cdots, c_n$;其**插值余项**为

$$R(x) = f(x) - P_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\prod_{i=0}^{n}(x - x_i),$$

式中 ξ 是区间中的某一点,这里假定被插函数 $f(x)$ 在区间上 $n+1$ 次可微.

(1) Lagrange 插值公式

引进 **Lagrange 多项式**

$$L_i(x) \equiv \prod_{\substack{l=0 \\ (l\neq i)}}^{n} \frac{x - x_l}{x_i - x_l},$$

显然有 $L_i(x_l) = \delta_{il}$.于是,插值问题的解可表达为

$$P_n(x) \equiv \sum_{i=0}^{n} y_i L_i(x) \equiv \sum_{i=0}^{n} y_i \prod_{\substack{l=0 \\ (l\neq i)}}^{n} \frac{x - x_l}{x_i - x_l}.$$

这是具有重要理论意义的 **Lagrange 插值公式**.

(2) Neville 插值算法

Lagrange 插值公式便于进行理论分析,也是导出数值微分或积分公式等的理

① W. E. Milne, *Numerical Calculus*, Princeton, 1949; § § 23, 24, 21.

② S. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, Springer, New York, 1980; p. 38.

论依据.但是,对于仅求某点之插值结果而言,实际计算则以某种**迭代算法**更为方便.

下面介绍 **Neville 插值算法**,特别适宜于计算机上采用,现列表并给出具体算法如表 13.1:

表 13.1

	$h=0$	1	2	3	…	
x_0	$y_0=N_{0,0}$					x_0-x
		$N_{1,1}$				
x_1	$y_1=N_{1,0}$		$N_{2,2}$			x_1-x
		$N_{2,1}$		$N_{3,3}$		
x_2	$y_2=N_{2,0}$		$N_{3,2}$			x_2-x
		$N_{3,1}$				
x_3	$y_3=N_{3,0}$					x_3-x

其中

$$N_{i,0}=y_i,\quad i=0,1,\cdots,n;$$

$$N_{i,k}=\frac{1}{x_i-x_{i-k}}\begin{vmatrix} N_{i-1,k-1} & x_{i-k}-x \\ N_{i,k-1} & x_i-x \end{vmatrix},\quad 1\leqslant k\leqslant i,\quad i=1,\cdots,n.$$

13.3.2 三次样条插值[①②]

在整个插值区间采用一个高次多项式进行插值,常会带来插值过程的不稳定性.另一方面,单纯的分段插值又不能保证良好的总体光滑度.**样条插值**是一种非局部性分段插值,能保证具有一定程度的总体光滑度.

三次样条插值的基本近似是在每一个子区间$[x_i,x_{i+1}]$将函数 $f(x)$都用三次多项式来代表:

$$S(x)=\alpha_i+\beta_i(x-x_i)+\frac{1}{2}\gamma_i(x-x_i)^2+\frac{1}{6}\delta_i(x-x_i)^3,$$

$$x_i\leqslant x\leqslant x_{i+1},\quad i=0,1,\cdots,n-1;$$

除要求在插值节点处满足插值条件

$$S(x_i)=f(x_i)=y_i\quad(i=0,1,\cdots,n)$$

外,还要求在每个分段点处满足直至二阶导数的连续性条件

$$S(x_i-0)=S(x_i+0),\quad S'(x_i-0)=S'(x_i+0),$$

$$S''(x_i-0)=S''(x_i+0)\equiv M_i\quad(i=1,2,\cdots,n-1);$$

---

① Stoer, Bulirsch, *Introduction to Numerical Analysis*; p. 93.

② 冯康等:《数值计算方法》,科学出版社,北京,1978;p. 13.

若再在两端点 x_0 和 x_n 各给定一个边界条件，插值问题即可唯一定解.

由此容易求得样条函数 $S(x)$ 中系数 $\alpha_i,\beta_i,\gamma_i,\delta_i(i=0,1,\cdots,n-1)$ 的下列结果：

$$\alpha_i = y_i,$$
$$\gamma_i = M_i,$$
$$\beta_i = \frac{y_{i+1}-y_i}{h_i} - \frac{h_i}{6}(M_{i+1}+2M_i),$$
$$\delta_i = \frac{1}{h_i}(M_{i+1}-M_i);$$

其中 $h_i=x_{i+1}-x_i$，而 y_i 和 M_i 则分别是节点 x_i 处的函数值（给定）和二阶导数值. 由 $S'(x_i-0)=S'(x_i+0)$ 条件可求得 $M_i(i=0,1,\cdots,n)$ 需满足下列方程：

$$\frac{h_{i-1}}{6}M_{i-1} + \frac{h_{i-1}+h_i}{3}M_i + \frac{h_i}{6}M_{i+1} = \frac{y_{i+1}-y_i}{h_i} - \frac{y_i-y_{i-1}}{h_{i-1}},$$
$$i=1,2,\cdots,n-1.$$

通常采用的边界条件有：

① 自然样条，$S''(x_0)=S''(x_n)=0$；附加方程为

$$M_0=0,\quad M_n=0;$$

② 给定导数值，$S'(x_0)=y_0'$，$S'(x_n)=y_n'$，可求得附加方程为

$$2h_0M_0 + h_0M_1 = 6\left[\frac{y_1-y_0}{h_0} - y_0'\right],$$
$$h_{n-1}M_{n-1} + 2h_{n-1}M_n = 6\left[y_n' - \frac{y_n-y_{n-1}}{h_{n-1}}\right];$$

③ 给定二阶导数值，$S''(x_0)=y_0''$，$S''(x_n)=y_n''$，附加方程为

$$M_0 = y_0'',\quad M_n = y_n''.$$

于是，结果导致关于 $M_0,M_1,\cdots,M_n$ 的下列线性方程组，

$$\begin{bmatrix} 2 & \lambda_0 & & & \\ \mu_1 & 2 & \lambda_1 & & \\ & \mu_2 & \ddots & \ddots & \\ & & \ddots & 2 & \lambda_{n-1} \\ & & & \mu_n & 2 \end{bmatrix} \begin{bmatrix} M_0 \\ M_1 \\ \vdots \\ M_{n-1} \\ M_n \end{bmatrix} = \begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix},$$

其中

$$\lambda_i = \frac{h_i}{h_{i-1}+h_i},\quad \mu_i = 1-\lambda_i = \frac{h_{i-1}}{h_{i-1}+h_i},\quad d_i = \frac{6}{h_{i-1}+h_i}\left\{\frac{y_{i+1}-y_i}{h_i} - \frac{y_i-y_{i-1}}{h_{i-1}}\right\},$$
$$i=1,2,\cdots,n-1;$$

而 $\lambda_0, d_0, \mu_n, d_n$ 则按不同边界条件由表 13.2 给出：

表 13.2

	①	②	③
λ_0	0	1	0
d_0	0	$\frac{6}{h_0}\left(\frac{y_1-y_0}{h_0}-y_0'\right)$	y_0''
μ_n	0	1	0
d_n	0	$\frac{6}{h_{n-1}}\left(y_n'-\frac{y_n-y_{n-1}}{h_{n-1}}\right)$	y_n''

13.4 数值微分和积分

数值微分或积分虽可以通过对插值多项式进行微分或积分而求得，但以采用**待定系数法**进行推导常常更为简便，而且较普遍适用.

13.4.1 数值微分①

设给定四点：$(x_0,y_0),(x_1,y_1),(x_2,y_2),(x_3,y_3)$，欲求 $y_1'=\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)_{x=x_1}$.

根据 Weierstrass 定理，可用多项式逼近为

$$y=c_0'+c_1'x+c_2'x^2+c_3'x^3,$$

$$\frac{\mathrm{d}y}{\mathrm{d}x}=c_1'+2c_2'x+3c_3'x^2;$$

于是，可得出下列方程组：

$$\begin{aligned}
-y_1'+0c_0'+c_1'+2c_2'x_1+3c_3'x_1^2&=0,\\
-y_0+c_0'+c_1'x_0+c_2'x_0^2+c_3'x_0^3&=0,\\
-y_1+c_0'+c_1'x_1+c_2'x_1^2+c_3'x_1^3&=0,\\
-y_2+c_0'+c_1'x_2+c_2'x_2^2+c_3'x_2^3&=0,\\
-y_3+c_0'+c_1'x_3+c_2'x_3^2+c_3'x_3^3&=0.
\end{aligned}$$

根据线性变换的定理，若以 $-1,c_0',c_1',c_2',c_3'$ 为五变量，则欲得非平凡解，即

$$(-1,c_0',c_1',c_2',c_3')\neq(0,0,0,0,0)$$

（现在 $-1\neq0$，上述情况满足），必须

① Milne, *Numerical Calculus*, §26, §27.

$$\begin{vmatrix} y_1' & 0 & 1 & 2x_1 & 3x_1^2 \\ y_0 & 1 & x_0 & x_0^2 & x_0^3 \\ y_1 & 1 & x_1 & x_1^2 & x_1^3 \\ y_2 & 1 & x_2 & x_2^2 & x_2^3 \\ y_3 & 1 & x_3 & x_3^2 & x_3^3 \end{vmatrix} = 0.$$

解出之，即得 y_1' 为 y_0, y_1, y_2, y_3 的线性组合.

另一方面，也可直接令

$$y_1' = c_0 y_0 + c_1 y_1 + c_2 y_2 + c_3 y_3$$

而确定其系数. 由于任何三次多项式都是 $1, (x-x_0), (x-x_0)^2, (x-x_0)^3$ 四函数的线性组合，因此，若 $y=1, (x-x_0), (x-x_0)^2, (x-x_0)^3$ 能满足上式，则任何三次多项式也就能满足上式了. 设 y 分别为此四函数，对于等距点 $x_i = x_0 + ih$ $(i=0,1,2,3)$ 容易求得系数 c_i 应满足：

$$\begin{aligned} 0 &= c_0 + c_1 + c_2 + c_3, \\ 1 &= hc_1 + 2hc_2 + 3hc_3, \\ 2h &= h^2 c_1 + 4h^2 c_2 + 9h^2 c_3, \\ 3h^2 &= h^3 c_1 + 8h^3 c_2 + 27h^3 c_3. \end{aligned}$$

解出 c_i 结果得到

$$y_1' = \frac{1}{6h}(-2y_0 - 3y_1 + 6y_2 - y_3) + \frac{h^3}{12} y^{(4)},$$

其中误差项是由插值余项求得：

$$\begin{aligned} R'(x_1) &= \frac{y^4(\xi)}{4!} \left\{ \frac{\mathrm{d}}{\mathrm{d}x} \prod_{i=0}^{3} (x - x_i) \right\}_{x=x_1} \\ &= \frac{y^{(4)}(\xi)}{24} (x_1 - x_0)(x_1 - x_2)(x_1 - x_3) \\ &= \frac{h^3}{12} y^{(4)}(\xi). \end{aligned}$$

还可导出其他一些微分公式，这里不予描述.

13.4.2　数值积分①②

数值积分也可采用待定系数法类似地进行推导. 这里先换用 Lagrange 插值公式

---

① Milne, *Numerical Calculus*; §§28—36.

② Stoer, Bulirsch, *Introduction to Numerical Analysis*; §3.1.

$$P_n(x) = \sum_{i=0}^{n} y_i L_i(x), \quad L_i(x) = \prod_{\substack{l=0 \\ (l \neq i)}}^{n} \frac{x - x_l}{x_i - x_l}$$

作为逼近函数来求 $\int_{x_0}^{x_n} y\mathrm{d}x$.

对于等距点 $x_i = x_0 + ih$ $(i=0,1,\cdots,n)$,引进 $x = x_0 + ht$,则有

$$L_i(x) = \lambda_i(t) \equiv \prod_{\substack{l=0 \\ (l \neq i)}}^{n} \frac{t-l}{i-l}.$$

于是,

$$\begin{aligned}
\int_{x_0}^{x_n} y\mathrm{d}x &= \int_{x_0}^{x_n} P_n(x)\mathrm{d}x \\
&= \sum_{i=0}^{n} y_i \int_{x_0}^{x_n} L_i(x)\mathrm{d}x \\
&= h\sum_{i=0}^{n} y_i \int_0^n \lambda_i(t)\mathrm{d}t = h\sum_{i=0}^{n} y_i \alpha_i, \\
\alpha_i &= \int_0^n \lambda_i(t)\mathrm{d}t.
\end{aligned}$$

例如,对 $n=2$,有

$$\begin{aligned}
\alpha_0 &= \int_0^2 \frac{(t-1)(t-2)}{(0-1)(0-2)}\mathrm{d}t = \frac{1}{2}\int_0^2 (t^2 - 3t + 2)\mathrm{d}t = \frac{1}{3}, \\
\alpha_1 &= \int_0^2 \frac{(t-0)(t-2)}{(1-0)(1-2)}\mathrm{d}t = -\int_0^2 (t^2 - 2t)\mathrm{d}t = \frac{4}{3}, \\
\alpha_2 &= \int_0^2 \frac{(t-0)(t-1)}{(2-0)(2-1)}\mathrm{d}t = \frac{1}{2}\int_0^2 (t^2 - t)\mathrm{d}t = \frac{1}{3};
\end{aligned}$$

于是,得到下列近似值:

$$\int_{x_0}^{x_2} y\mathrm{d}x = \frac{h}{3}(y_0 + 4y_1 + y_2) - \frac{y^{(4)}}{90}h^5,$$

这是著名的 **Simpson 法则**. 关于所引误差项,这里不作论证.

几个常用的类似数值积分法则列于表 13.3(注意到 $\alpha_i = \frac{1}{s}\sigma_i = \frac{n}{ns}\sigma_i$).

表 13.3

n	$\sigma_i = s\alpha_i$	ns	误差	名称
1	1 1	2	$h^3\ \frac{1}{12}y^{(2)}$	梯形法则
2	1 4 1	6	$h^5\ \frac{1}{90}y^{(4)}$	Simpson 法则
3	1 3 3 1	8	$h^5\ \frac{3}{80}y^{(4)}$	3/8 法则
4	7 32 12 32 7	90	$h^7\ \frac{8}{945}y^{(6)}$	Milne 法则

以上这些结果都属于

$$\int_{x_0}^{x_n} y\mathrm{d}x = A_0 y_0 + A_1 y_1 + \cdots + A_n y_n$$

这种类型，它包含两端点纵坐标在内，统称之为 **Newton-Cotes 闭型积分公式**.

另有一类 **Newton-Cotes 开型积分公式**，属于

$$\int_{x_0}^{x_n} y\mathrm{d}x = A_1 y_1 + A_2 y_2 + \cdots + A_{n-1} y_{n-1}$$

这种类型，它不含端点纵坐标. 现以

$$\int_{x_0}^{x_4} y\mathrm{d}x = \frac{4h}{3}(2y_1 - y_2 + 2y_3) + \frac{28y^{(4)}h^5}{90}$$

为例，用待定系数法推导如下.

分别以 $y=1,(x-x_0),(x-x_0)^2$ 代入前式后得到

$$4h = A_1 + A_2 + A_3,$$
$$8h = A_1 + 2A_2 + 3A_3,$$
$$\frac{64}{3}h = A_1 + 4A_2 + 9A_3;$$

解之得 $A_1=\frac{8}{3}h, A_2=-\frac{4}{3}h, A_3=\frac{8}{3}h$，此即引例之结果. 至于所引误差项，则需另行计算，这里不予讨论.

待定系数法还可用于推导更普遍类型的公式. 例如，设

$$\int_{x_0}^{x_2} y\mathrm{d}x = A_{-1}y_{-1} + A_0 y_0 + A_1 y_1 + By_0'.$$

分别以 $y=1,(x-x_0),(x-x_0)^2,(x-x_0)^3$ 代入后得到

$$2h = A_{-1} + A_0 + A_1,$$
$$2h = -A_{-1} + A_1 + \frac{1}{h}B,$$
$$\frac{8}{3}h = A_{-1} + A_1,$$
$$4h = -A_{-1} + A_1;$$

解之得 $A_{-1}=-\frac{2}{3}h, A_0=-\frac{2}{3}h, A_1=\frac{10}{3}h, B=-2h^2$. 结果有

$$\int_{x_0}^{x_2} y\mathrm{d}x = \frac{2h}{3}(5y_1 - y_0 - y_{-1}) - 2h^2 y_0' + \frac{7}{45}h^5 y^{(5)},$$

其中引用了误差项而不作推导.

13.5　微分方程的数值解法

由于现代科学技术实践的需要与计算机的不断更新换代，求解微分方程的有

限差分法得到很大进展，而且还发展出了崭新的**有限元法**，目前都有广泛应用.

下面仅对常微分方程初值问题的数值解法作一简要介绍.

13.5.1 Runge-Kutta 法①

对于一阶常微分方程

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y' = f(x,y),$$

在区间(x_p, x_{p+1})进行积分可得

$$y_{p+1} = y_p + \int_{x_p}^{x_{p+1}} f(x,y)\mathrm{d}x.$$

由于对 $f(x,y)$采取不同的逼近式，可得不同的公式.

例如，若分别用两端值、平均值或中点值代替之，则得 **Euler 折线法**的结果：

$y_{p+1} = y_p + hf_p$，向前差公式，

$y_{p+1} = y_p + hf_{p+1}$，向后差公式，

$y_{p+1} = y_p + \frac{1}{2}h(f_{p+1} + f_p)$，平均(梯形)公式，

$y_{p+1} = y_p + hf_{p+\frac{1}{2}}$，中心差公式；

最后一个常改为

$$y_{p+1} = y_{p-1} + 2hf_p.$$

注意到，这里采用等距点 $h = x_{p+1} - x_p$，$f_l = f(x_l, y_l)$.

若将 Euler 中心差公式中的 $y_{p+\frac{1}{2}}$ 在 x_p 处展开为

$$y_{p+\frac{1}{2}} \approx y_p + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)_{x=x_p} \frac{h}{2} = y_p + \frac{1}{2}hf(x_p, y_p),$$

则得**二阶 Runge-Kutta 算法**：

$$y_{p+1} = y_p + k_2,$$

$$k_2 = hf\left(x_p + \frac{1}{2}h, y_p + \frac{1}{2}k_1\right), \quad k_1 = hf(x_p, y_p).$$

采用类似结构和作法，可得**四阶 Runge-Kutta 算法**如下：

$$y_{p+1} = y_p + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

$$k_1 = hf(x_p, y_p),$$

$$k_2 = hf\left(x_p + \frac{1}{2}h, y_p + \frac{1}{2}k_1\right),$$

$$k_3 = hf\left(x_p + \frac{1}{2}h, y_p + \frac{1}{2}k_2\right),$$

---

① 冯康等：《数值计算方法》；§ 11.3.1，§ 11.8.

$$k_4 = hf(x_p + h, y_p + k_3).$$

由此公式，若已知 y_p，就可依次算得 k_1 至 k_4，从而求得 y_{p+1}. 因此，对于初值问题，由给定的 y_0，即可通过令 $p=0,1,2,\cdots$ 逐步求得 $y_1, y_2, y_3, \cdots$ 这样继续进行下去. 它具有四阶精度，是广为应用的一种算法.

应该注意，Runge-Kutta 算法是非线性单步差分格式.

顺便指出，本小节的方法同样适用于一阶常微分方程组，只要将所有 y, k 和 f 相应地都换成矢量 $\boldsymbol{y}, \boldsymbol{k}$ 和 $\boldsymbol{f}$ 即可.

13.5.2 Adams 法①②

对于上一小节的公式

$$y_{p+1} = y_p + \int_{x_p}^{x_{p+1}} f(x,y)\mathrm{d}x,$$

将其中的 $f(x,y)$ 用 Lagrange 插值公式予以逼近并进行积分，可得 **Adams 法**的结果. 由于有外插与内插之别，以及多项式次数的多寡，而有下列不同之公式(精度为 $q+1$ 阶).

(1) Adams-Bashforth 公式

采用外插方式的 **Adams-Bashforth 公式**形式为

$$y_{p+1} = y_p + h\sum_{i=0}^{q} \beta_{qi} f_{p-i},$$

$$\beta_{qi} = \int_0^1 \prod_{\substack{l=0 \\ (l\neq i)}}^{q} \frac{t+l}{-i+l}\mathrm{d}t, \quad i = 0,1,\cdots,q;$$

系数 β_{qi} 见表 13.4. 因为公式右端不含 y_{p+1}，称为**显式**. 可以看出，$q=0$ 时即 Euler 向前差公式.

表 13.4

β_{qi} \ i	0	1	2	3
β_{0i}	1			
$2\beta_{1i}$	3	-1		
$12\beta_{2i}$	23	-16	5	
$24\beta_{3i}$	55	-59	37	-9

(2) Adams-Moulton 公式

采用内插方式的 **Adams-Moulton 公式**形式为

① Stoer, Bulirsch, *Introduction to Numerical Analysis*; §7.2.6.

② 冯康等:《数值计算方法》; §11.6.

$$y_{p+1}=y_p+h\sum_{i=0}^{q}\beta_{qi}f_{p+1-i},$$

$$\beta_{qi}=\int_0^1\prod_{\substack{l=0\\(l\neq i)}}^{q}\frac{t-l}{i-l}\mathrm{d}t,\quad i=0,1,\cdots,q;$$

系数 β_{qi} 见表 13.5. 因为公式右端隐含有 y_{p+1}，称为**隐式**. 可以看出，$q=0,1$ 时分别对应于 Euler 向后差及平均公式.

表 13.5

β_{qi} \ i	0	1	2	3
β_{0i}	1			
$2\beta_{1i}$	1	1		
$12\beta_{2i}$	5	8	-1	
$24\beta_{3i}$	9	19	-5	1

应该指出，Adams 法须用$(q+1)$个结点上的值来计算 y_{p+1}，是**多步法**(当 $q>1$ 时)，须用**起步算法**. Adams 法采用的是线性多步差分格式.

13.5.3 预估校正法[①]

实践中常采用一个显式公式作为**预估式**，然后用一个隐式公式作为**校正式**，称为**预估校正法**；其中还需要某种**起步算法**.

例如，采用 Newton-Cotes 开型积分公式

$$y_{p+1}=y_{p-3}+\frac{4}{3}h(2y_p'-y_{p-1}'+2y_{p-2}')+\frac{28}{90}h^{(5)}y^{(5)}$$

作为预估式，Simpson 公式

$$y_{p+1}=y_{p-1}+\frac{1}{3}h(y_{p+1}'+4y_p'+y_{p-1}')-\frac{1}{90}h^5y^{(5)}$$

作为校正式.

为了起步，除初值 y_0 外，还需 y_{-1},y_1,y_2；可由下列**起步式**：

$$y_1=y_0+\frac{1}{24}h(7y_1'+16y_0'+y_{-1}')+\frac{1}{4}h^2y_0''-\frac{1}{180}h^5y^{(5)},$$

$$y_{-1}=y_0-\frac{1}{24}h(y_1'+16y_0'+7y_{-1}')+\frac{1}{4}h^2y_0''+\frac{1}{180}h^5y^{(5)},$$

$$y_2=y_0+\frac{2}{3}h(5y_1'-y_0'-y_{-1}')-2h^2y_0''+\frac{7}{45}h^5y^{(5)}$$

① Milne, *Numerical Calculus*; § 38.

来提供. 首先由 $y_1'=y_0'+hy_0''$ 和 $y_{-1}'=y_0'-hy_0''$ 给出 y_1' 和 y_{-1}' 的试探值，然后可由前二式并且结合微分方程及再求导所得方程经反复迭代以求得 y_{-1} 和 y_1 的起步值. 最后一式则给出 y_2 的起步值. 这就是此法的起步算法.

有了这些起步值，就可由所给公式求得下一点 y_3 的预估值和校正值. 由误差项可知

$$\text{校正值之误差约为}\frac{1}{29}\times(\text{校正值}-\text{预估值}).$$

若此值符合精度要求，即可继续进行下去. 否则需缩短步长 h，将步长减半约使误差减为 1/32. 若两者之差可忽略不计，可试将步长加倍，若仍有足够精度，可使工作量减半.

下列 Adams 公式：

$$y_{p+1}=y_p+\frac{h}{12}(23y_p'-16y_{p-1}'+5y_{p-2}'),$$

$$y_{p+1}=y_p+\frac{h}{24}(9y_{p+1}'+19y_p'-5y_{p-1}'+y_{p-2}'),$$

也是常用的具有四阶精度的预估校正式.

13.5.4　二阶常微分方程

对于二阶常微分方程的初值问题：

$$y''=f(x,y,y'),\quad y(x_0)=y_0,\quad y'(x_0)=y_0',$$

同样可用预估校正法求解.

例如，y_{p+1}' 的值由

$$y_{p+1}'=y_{p-3}'+\frac{4}{3}h(2y_p''-y_{p-1}''+2y_{p-2}'')$$

预估，y_{p+1} 的值可由

$$y_{p+1}=y_{p-1}+\frac{1}{3}h(y_{p+1}'+4y_p'+y_{p-1}')$$

求得，而 y'' 的值则由微分方程求得；然后用

$$y_{p+1}'=y_{p-1}'+\frac{1}{3}h(y_{p+1}''-4y_p''+y_{p-1}'')$$

校正 y_{p+1}'. 若校正量相当大，则需迭代校正之. 至于起步算法，用上一小节的起步式来实现，只是把每个公式里的 y,y',y'' 换成 y',y'',y'''.

对于不含一阶导数的情形：

$$y''=f(x,y),$$

可分别采用

$$y_{p+1}=y_p+y_{p-2}-y_{p-3}+\frac{1}{4}h^2(5y_p''+2y_{p-1}''+5y_{p-2}'')+\frac{17}{240}h^6y^{(6)},$$

$$y_{p+1}=2y_p-y_{p-1}+\frac{1}{12}h^2(y''_{p+1}+10y''_p+y''_{p-1})-\frac{1}{241}h^6y^{(6)}$$

作为预估式和校正式;起步算法与上面相类似.

13.6 数值计算方法程序库

关于数值计算方法,包括本书(第 4、13 章及其他处)讨论过的,以及大量未提及的,目前大都有现成的标准的**数值计算方法程序库**可供调用,甚至可从互联网上下载,没有必要自己去编程序.

文献《计算物理》[①]有专门一章介绍如何应用**科学程序库**和**万维网**来获取所需程序.

一些可资用的标准科学和数学库包括:

LAPACK:线性代数程序包
SLATEC:综合数学和统计程序包
Netlib:免费数学程序库的万维网衍生库
IMSL:国际数学与统计程序库
ESSL:工程和科学子程序库(IBM)
DXML:高级数学程序库(DEC)
NAG:数值算法组(UK Labs)
BLAS:基本线性代数子程序(积木)

其内容涵盖有:

线性代数演算	矩阵运算
线性方程组的解	本征系统分析
插值、拟合	微分方程
根、零点、极值	随机数演算
统计函数	数值求积

该书并介绍了获取适用子程序的信息(§§15.8,15.9),最后三节(§§15.12—15.14)还列出了 LAPACK,Netlib 和 SLATEC 的简短目录.

① R. H. Landau, M. J. Paéz, *Computational Physics, Problem Solving with Computers*, Wiley and Sons, New York, 1997; §§15.8, 15.9, 15.12—15.14.

索　引

D

E

F

G

重排后记

本书自2001年5月出版后，获得物理学界的普遍好评，2009年本书又纳入《中外物理学精品书系》，获国家出版基金资助，排为16开本重新出版.

因彭桓武先生2009年已驾鹤西去，重排时徐锡申先生再次认真地审阅了书稿，改正了书中的一些遗误.

数理物理是物理现象数学模型的理论，由于本书书名国内少见，重排时，按徐锡申先生的提议，特意添加了副书名:《物理需用线性高等数学导引》，以期使更广大读者对其基本内容能够一目了然.

本书作为《中外物理学精品书系·经典系列》之一重排出版，以此告慰"两弹一星元勋"彭桓武先生.

北京大学出版社

2012年7月